AFRL-OSR-VA-TR-2015-0137

PREDICTING AND CONTROLLING COMPLEX NETWORKS

Ying Cheng Lai
ARIZONA STATE UNIVERSITY

06/22/2015
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTC
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
26-06-2015

2. REPORT TYPE
Final Performance

3. DATES COVERED *(From - To)*
01-04-2010 to 31-03-2015

4. TITLE AND SUBTITLE
PREDICTING AND CONTROLLING COMPLEX NETWORKS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-10-1-0083

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ying Cheng Lai

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ARIZONA STATE UNIVERSITY
660 S MILL AVE STE 312
TEMPE, AZ 85281 US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The principal Objective of the project was to develop methods to predict and control complex networks. For prediction, a number of methods were articulated and tested to uncover the structures and topologies of complex networks as well as various dynamical processes on the networks based solely on time series data or measured signals. A compressive sensing based framework for network and nonlinear dynamical systems reconstruction was pioneered. For control, key issues including linear controllability of complex networks, control energy, control of collective dynamics, and control of nonlinear dynamics on complex networks were addressed. A number of new phenomena in complex dynamical systems were uncovered and understood, and computational paradigms were established for prediction and control.

15. SUBJECT TERMS
PREDICTING, CONTROLLING

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ying Cheng Lai
U	U	U	UU		19b. TELEPHONE NUMBER *(Include area code)* 480-965-6668

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Final Report

This Final Report summarizes activities under the Air Force Office of Scientific Research (AFOSR) Grant No. FA9550-10-1-0083 entitled "PREDICTING AND CONTROLLING COMPLEX NETWORKS" from 1 April 2010 to 31 March 2015. PI is Ying-Cheng Lai from Arizona State University (ASU).

Contents

1 Objectives **2**

2 List of Publications **2**

3 Accomplishments and New Findings **5**

3.1 Uncovering the full topology of oscillator networks 5

3.2 Cascading failures and the emergence of cooperation in evolutionary-game based models of social and economic networks 6

3.3 Information explosion on complex networks and control 7

3.4 Pattern formation, synchronization and outbreak of biodiversity in cyclically competing games 8

3.5 Predicting catastrophes in nonlinear dynamical systems by compressive sensing 9

3.6 Time-series based prediction of complex oscillator networks via compressive sensing 11

3.7 Reconstruction of social networks based on evolutionary-game data via compressive sensing 12

3.8 Optimizing controllability of complex networks by minimum structural perturbations 15

3.9 Detecting hidden nodes in complex networks from time series based on compressive sensing 17

3.10 Forecasting synchronizability of complex networks from data 20

3.11 Emergence of grouping in multi-resource minority game dynamics 21

3.12 Optimizing cooperation on complex networks in the presence of failure 23

3.13 Exact controllability of complex networks 24

3.14 Emergence of scaling in human-interest dynamics 26

3.15 Robustness of chimera states in complex dynamical systems 28

3.16 Uncovering hidden nodes in complex networks in the presence of noise 30

3.17 Reconstructing propagation networks with natural diversity and identifying hidden source . . 32

3.18 Spatiotemporal patterns and predictability of cyberattacks 33

4 Personnel Supported and Theses Supervised by PI **35**

4.1 Personnel Supported 35

4.2 Theses supervised by PI in the project area 35

5 Interactions/Transitions **35**

5.1 Collaboration with AFOSR scientist 35

5.2 Invited talks on topics derived from the project 35

6 Past Honors **37**

1 Objectives

- **Prediction**: to develop methods to uncover the structures and topologies of complex networks as well as various dynamical processes from time series or data.
- **Control**: to understand the controllability of complex networks and articulate methods to control dynamical processes that they support.

2 List of Publications

1. J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Noise bridges dynamical correlation and topology in coupled oscillator networks," *Physical Review Letters* **104**, 058701, 1-4 (2010).
2. H.-J. Shi, R. Yang, W.-X. Wang, and Y.-C. Lai, "Basins of attraction for species extinction and coexistence in spatial rock-paper-scissors games," *Physical Review E (Rapid Communications)* **81**, 030901(R), 1-4 (2010).
3. W.-X. Wang, Y.-C. Lai, and C. Grebogi, "Effect of epidemic spreading on species coexistence in spatial games," *Physical Review E* **81**, 046113, 1-4 (2010). One figure from this paper was selected for "Kaleidoscope" of PRE.
4. R. Yang, W.-X. Wang, Y.-C. Lai, and C. Grebogi, "Role of intraspecific competition in the coexistence of mobile populations in spatially extended ecosystems," *Chaos* **20**, 023113, 1-6 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the June 1, 2010 issue (http://www.vjbio.org).
5. X. Ni, W.-X. Wang, Y.-C. Lai, and C. Grebogi, "Cyclic competition of mobile species on continuous space: pattern formation and coexistence," *Physical Review E* **82**, 066211, 1-8 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the January 1, 2011 issue (http://www.vjbio.org).
6. X. Ni, R. Yang, W.-X. Wang, Y.-C. Lai, and C. Grebogi, "Basins of coexistence and extinction in spatially extended ecosystems of cyclically competing species," *Chaos* **20**, 045116, 1-8 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the January 1, 2011 issue (http://www.vjbio.org).
7. W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, "Predicting catastrophes in nonlinear dynamical systems by compressive sensing," *Physical Review Letters* **106**, 154101, 1-4 (2011).
8. W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, "Pattern formation, synchronization and outbreak of biodiversity in cyclically competing games," *Physical Review E* **83**, 011917, 1-9 (2011).
9. R.-R. Liu, W.-X. Wang, Y.-C. Lai, G.-R. Chen, and B.-H. Wang, "Optimal convergence in naming game with geography-based negotiation on small-world networks," *Physics Letters A* **375**, 363-367 (2011).
10. H.-X. Yang, W.-X. Wang, Y.-B. Xie, Y.-C. Lai, and B.-W. Wang, "Transportation dynamics on networks of mobile agents," *Physical Review E* **83**, 016102, 1-5 (2011).
11. L.-L. Jiang, M. Perc, W.-X. Wang, Y.-C. Lai and B.-H. Wang, "Impact of link deletions on public cooperation in scale-free networks," *Europhysics Letters* **93**, 40001, 1-6 (2011).
12. W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and M. A. F. Harrison, "Time-series based prediction of complex oscillator networks via compressive sensing," *Europhysics Letters* **94**, 48006, 1-6 (2011).

13. L. Huang and Y.-C. Lai, "Cascading dynamics in complex quantum networks," *Chaos* **21**, 025107, 1-6 (2011). This work was selected by July 2011 issue of Virtual Journal of Quantum Information (http://www.vjquantuminfo.org).
14. W.-X. Wang, Y.-C. Lai, and D. Armbruster, "Cascading failures and the emergence of cooperation in evolutionary game based models of social and economical networks," *Chaos* **21**, 033112, 1-12 (2011).
15. H.-X. Yang, W.-X. Wang, Y.-C. Lai, Y.-B. Xie, and B.-H. Wang, "Control of epidemic spreading on complex networks by local traffic dynamics," *Physical Review E (Rapid Communication)* **84**, 045101(R), 1-4 (2011).
16. W.-X. Wang, Y.-C. Lai, C. Grebogi, and J.-P. Ye, "Network reconstruction based on evolutionary-game data via compressive sensing," *Physical Review X* **1**, 021021, 1-7 (2011).
17. R.-R. Liu, W.-X. Wang, Y.-C. Lai, and B.-H. Wang, "Cascading dynamics on random networks: crossover in phase transition," *Physical Review E* **85**, 026110, 1-5 (2012).
18. W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, "Optimizing controllability of complex networks by small structural perturbations," *Physical Review E* **85**, 026115, 1-5 (2012).
19. G.-M. Zhu, H.-J. Yang, R. Yang, J. Ren, B. Li, and Y.-C. Lai, "Uncovering evolutionary ages of nodes in complex networks," *European Journal of Physics B* **85**, 106, 1-6 (2012).
20. G. Yan, J. Ren, Y.-C. Lai, C. H. Lai, and B. Li, "Controlling complex networks - how much energy is needed?" *Physical Review Letters* **108**, 218703, 1-5 (2012).
21. R.-Q. Su, X. Ni, W.-X. Wang, and Y.-C. Lai, "Forecasting synchronizability of complex networks from data," *Physical Review E* **85**, 056220, 1-11 (2012).
22. H.-X. Yang, W.-X. Wang, Y.-C. Lai, and B.-H. Wang, "Traffic-driven epidemic spreading on networks of mobile agents," *Europhysics Letters* **98**, 68003, 1-5 (2012).
23. L.-L. Jiang, W.-X. Wang, Y.-C. Lai, and X. Ni, "Multi-armed spirals and multi-pairs antispirals in spatial rock-paper-scissors games," *Physics Letters A* **376**, 2292-2297 (2012).
24. R.-Q. Su, W.-X. Wang, and Y.-C. Lai, "Detecting hidden nodes in complex networks from time series," *Physical Review E (Rapid Communication)* **85**, 065201(R), 1-4 (2012).
25. R. Yang, Y.-C. Lai, and C. Grebogi, "Forecasting the future: is it possible for time-varying nonlinear dynamical systems?" *Chaos* **22**, 033119, 1-6 (2012).
26. F. Ricci, R. Tonelli, L. Huang, and Y.-C. Lai, "Onset of chaotic phase synchronization in complex networks of coupled heterogeneous oscillators," *Physical Review E* **86**, 027201, 1-4 (2012).
27. W.-X. Wang, J. Ren, Y.-C. Lai, and B. Li, "Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series," *Chaos* **22**, 033131, 1-8 (2012).
28. Z.-G. Huang, J.-Q. Zhang, J.-W. Dong, L. Huang, and Y.-C. Lai, "Emergence of grouping in multi-resource minority game dynamics," *Nature Scientific Reports* **2**, 703, 1-8 (2012).
29. Y.-Z. Chen and Y.-C. Lai, "Optimizing cooperation on complex networks in the presence of failure," *Physical Review E (Rapid Communications)* **86**, 045101(R), 1-4 (2012).
30. L. Huang, Y.-C. Lai, and M. A. F. Harrison, "Probing complex networks from measured time series," *International Journal of Bifurcation and Chaos* **22**, 1250236, 1-12 (2012).
31. H.-X. Yang, W.-X. Wang, and Y.-C. Lai, "Traffic-driven epidemic outbreak on complex networks: how long does it take?" *Chaos* **22**, 043146, 1-5 (2012).

32. Z. Zhou, Z.-G. Huang, L. Huang, Y.-C. Lai, L. Yang, and D.-S. Xue, "Universality of flux-fluctuation law in complex dynamical systems," *Physical Review E* **87**, 012808, 1-6 (2013).

33. J.-Q. Zhang, Z.-G. Huang, J.-Q. Dong, L. Huang, and Y.-C. Lai, "Controlling collective dynamics in complex, minority-game resource-allocation systems," *Physical Review E* **87**, 052808, 1-9 (2013).

34. J.-P. Park, Y.-H. Do, Z.-G. Huang, and Y.-C. Lai, "Persistent coexistence of cyclically competing species in spatially extended ecosystems," *Chaos* **23**, 023128, 1-9 (2013).

35. Z.-Z. Yuan, C. Zhao, Z.-R. Di, W.-X. Wang, and Y.-C. Lai, "Exact controllability of complex networks," *Nature Communications* **4**, 2447, 1-9 (2013).

36. B.-S. Kim, Y.-H. Do, and Y.-C. Lai, "Emergence of synchronization and size scaling in moving-agent networks," *Physical Review E* **88**, 042818, 1-7 (2013).

37. Z.-D. Zhao, Z.-M. Yang, Z.-K. Zhang, T. Zhou, Z.-G. Huang, and Y.-C. Lai, "Emergence of scaling in human-interest dynamics," *Nature Scientific Reports* **3**, 3472, 1-7 (2013).

38. K. Gong, M. Tang, P. M. Hui, Y. Do, and Y.-C. Lai, "An efficient immunization strategy for community networks," *PLoS One* **8**, e83489, 1-11 (2013).

39. N. Yao, Z.-G. Huang, Y.-C. Lai, and Z.-G. Zheng, "Robustness of chimera states in complex dynamical systems," *Nature Scientific Reports* **3**, 3522, 1-8 (2013).

40. R.-Q. Su, Y.-C. Lai, X. Wang, and Y.-H. Do, "Uncovering hidden nodes in complex network in the presence of noise," *Nature Scientific Reports* **4**, 3944, 1-7 (2014).

41. W. Wang, M. Tang, H. Yang, Y.-H. Do, Y.-C. Lai, and G.-W. Lee, "Asymmetrically interacting spreading dynamics on complex layered networks," *Nature Scientific Reports* **4**, 5097, 1-8 (2014).

42. Z.-S. Shen, W.-X. Wang, Y. Fan, Z.-R. Di, and Y.-C. Lai, "Reconstructing propagation networks with natural diversity and identifying hidden source," *Nature Communications* **5**, 4323, 1-10 (2014).

43. H.-F. Zhang, Z.-X. Wu, M. Tang, and Y.-C. Lai, "Effects of behavioral response and vaccination policy on epidemic spreading - an approach based on evolutionary-game dynamics," *Nature Scientific Reports* **4**, 5666, 1-10 (2014).

44. R.-Q. Su, Y.-C. Lai, and X. Wang, "Identifying chaotic FitzHugh-Nagumo neurons using compressive sensing," *Entropy* **16**, 3889-3902 (2014).

45. Y.-C. Lai, "Controlling complex, nonlinear dynamical networks," *National Science Review* **1**, 339-341 (2014).

46. Y.-Z. Chen, Z.-G. Huang, and Y.-C. Lai, "Controlling extreme events on complex networks," *Nature Scientific Reports* **4**, 6121, 1-10 (2014).

47. H.-F. Zhang, J.-R. Xie, M. Tang, and Y.-C. Lai, "Suppression of epidemic spreading in complex networks by local information based behavioral responses," *Chaos* **24**, 043106, 1-7 (2014).

48. Z.-Z. Yuan, C. Zhao, W.-X. Wang, Z.-R. Di, and Y.-C. Lai, "Exact controllability of multiplex networks," *New Journal of Physics* **16**, 103036, 1-24 (2014).

49. L.-Z. Wang, Z.-G. Huang, Z.-H. Rong, X.-F. Wang, and Y.-C. Lai, "Emergence and evolution of online social networks," *PLoS ONE* **9**(11), e111013, 1-6 (2014).

50. Y.-Z. Chen, Z.-G. Huang, S.-H. Xu, and Y.-C. Lai, "Spatiotemporal patterns and predictability of cyberattacks," *PLoS ONE*, accepted (to appear in June 2015).

3 Accomplishments and New Findings

3.1 Uncovering the full topology of oscillator networks

Previous efforts in network science and engineering were mostly focused on network structures and their effects on various dynamical processes taking place on the network. The types of processes investigated include synchronization, virus spreading, traffic flow, and cascading failures. A typical approach in the field was to implement a particular dynamical process of interest on networks whose connecting topologies are completely specified. Often, real-world networks such as the Internet, the power grids, transportation networks, and various biological and social networks were used as examples to demonstrate the relevance of the dynamical phenomena found from model networks. While this line of research was necessary for discovering and understanding various fundamental phenomena in complex networks, the "inverse" problem of network prediction is also extremely important. The basic hypothesis underlying the inverse problem is that the detailed structure of the network and the node dynamics are totally unknown, but only a limited set of signals or time series measured from the network is available. *The question is whether the intrinsic structure of the network can be inferred solely from the set of measured time series.* Compared with the "direct" network-dynamics problem, the inverse problem received relatively little attention due to the extremely challenging nature of the problem. Nonetheless it is of paramount importance to address the problem not only for advancing network science and engineering, but also for meeting the need to address an array of realistic applications where large-scale, complex networks arise.

During the performance period, we developed a framework to uncover the full topology of oscillatory networks from time series in the presence of noise and time delay. The reason to consider noise and time delay is that they are ubiquitous in real-world complex systems. We obtained the surprising result that, in the presence of noise, it becomes generally possible to precisely identify interactions based *solely* on the correlations among measured time series from various nodes in the underlying network. In particular, by defining the dynamical correlation between pairwise oscillators as the product of their state differences from the respective time-averaged values, we obtained a dynamical correlation matrix that can be calculated purely from time series. Analytically, we found the existence of a one-to-one correspondence between the correlation matrix and the network connection matrix in the presence of noise. In this sense, it can be said that *noise bridges dynamics and topology*, facilitating inference of network structures. Indeed, understanding the effects of noise on dynamical systems has been a fundamental issue in nonlinear and statistical physics. While there had been previous works on the interplay between the collective dynamics and the topology of complex systems under noise (including our own previous AFOSR-sponsored work on predicting node degrees of complex networks), taking advantage of noise to predict the *full connecting topology* of unknown complex oscillatory networks had not been achieved prior to our work. Our philosophy is in fact to *make use of noise to obtain knowledge about the network from noisy data*.

Mathematically, our main result and its applicable setting can be summarized, as follows. Consider a network of N coupled oscillators. Each oscillator, when decoupled, satisfies $\dot{\mathbf{x}}_i = \mathbf{F}_i[\mathbf{x}_i]$, where $\mathbf{x}_i$ denotes the d-dimensional state variable. The dynamics of the whole time-delayed system in a noisy environment can be described as

$$\dot{\mathbf{x}}_i(t) = \mathbf{F}_i[\mathbf{x}_i(t)] - c\sum_{j=1}^{N} L_{ij}\mathbf{H}[\mathbf{x}_j(t-\tau)] + \eta_i, \tag{1}$$

where $\mathbf{H}$ denotes the coupling function, L_{ij} is the element of Laplacian matrix of the underlying network, c denotes the coupling strength, τ is the time delay, and η_i is a Gaussian noisy process of zero mean and variance σ^2. The element of the dynamical correlation matrix (between time series from nodes i and j) is

$C_{ij} = \langle \xi_i(t)\xi_j(t) \rangle$, where $\xi_i(t) \equiv x_i(t) - (1/N)\sum_{i=1}^{N} x_i(t)$. Our main result was the following formula:

$$\frac{\sigma^2}{2c} C^{\dagger}_{ij} \approx \begin{cases} L_{ij} + c\tau(k_i + k_j), & \text{if } i \text{ connects with } j \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where $C^{\dagger}_{ij}$ is the element of the pseudo-inverse of the correlation matrix, k_i and k_j are the degrees of nodes i and j, respectively. This formula indicates that the network structure can be inferred through the off-diagonal elements $C^{\dagger}_{ij}$ of the dynamical correlation matrix based solely on the measured time series. After $\hat{\mathbf{L}}$ is predicted, the time delay τ can be estimated as

$$\tau \approx \left\langle \frac{\left[\hat{\mathbf{L}} - \frac{\sigma^2}{2c}\hat{\mathbf{C}}^{\dagger}\right]_{ij}}{c[\hat{\mathbf{L}}^2]_{ij}} \right\rangle_{i \neq j, L_{ij} \neq 0, (\hat{\mathbf{L}}^2)_{ij} \neq 0}, \tag{3}$$

where the subscript in the average $\langle \cdot \rangle$ covers all possible pairs of i and j by excluding the diagonal elements in the matrices $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}^2$, and all pairs with zero elements in the matrix $\hat{\mathbf{L}}$ or $\hat{\mathbf{L}}^2$. Excluding zero elements can effectively reduce the estimation error for τ.

A detailed account of the mathematical theory and results from extensive numerical tests were summarized in the following two papers:

1. J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Noise bridges dynamical correlation and topology in coupled oscillator networks," *Physical Review Letters* **104**, 058701(1-4) (2010).
2. W.-X. Wang, J. Ren, Y.-C. Lai, and B. Li, "Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series," *Chaos* **22**, 033131, 1-8 (2012).

3.2 Cascading failures and the emergence of cooperation in evolutionary-game based models of social and economic networks

A hallmark of the economic recession in 2008 was the collapse and bankruptcy of a large number of financial institutions and corporations on a scale that has not been seen since the great depression. The manner by which the failures occur may be described as a cascading process, where the initial collapse of one or a few institutions, for example, triggered the failures of many others. While sophisticated economic and social models can be articulated to describe the process of cascading collapses, to obtain fundamental insights we were interested in a "minimal" model that can capture the major generic ingredients of the process, which are independent of the system details. It was hoped that the model could then lead to insights into the prevention of such cascades. We constructed such a model, carried out a detailed theoretical analysis and extensive numerical tests, and explored implications.

Our model was based on evolutionary games on large networks, a powerful paradigm to study a variety of self-organized behaviors in natural, social, and economic systems. In previous works, the tolerance of individuals to elimination or death had not been investigated in a comprehensive manner, but such a process is important in different areas. For example, consider the bankruptcy of agents in economic systems. For any agent, a lowest amount of profit should be maintained for it to survive, which comes from the interactions with other agents in a certain time period for continuous investment into the future. Another example is ecosystems, where individuals compete and cooperate for essential life-sustaining resources. If some minimal requirement for resources cannot be satisfied, individuals will die. In our work, we incorporated an elimination mechanism into the gaming rules to better mimic the evolution of cooperative behavior in realistic systems. In particular, we assigned a tolerance parameter to every individual in the network, which is the lowest payoff needed for an individual to survive. Taking into account the diversity in real-world systems, each individual can have its own tolerance. For example, the number of interactions of an individual is a characteristic to distinguish it from others, so it can be used to define the tolerance. In a network, the death of an individual leads to removal of the corresponding node together with all the links with the other nodes. The game and network thus co-evolve as a result of the elimination. Our main finding was that, in the presence of defectors, a cascading process of death of individuals can occur in relatively short time, which can even spread to the whole network, leading to complete extinction. Strikingly, we found that a pure cooperation state can emerge after the cascade terminates, in which the exclusive survivors are cooperators. This phenomenon occurs regardless of the type of games and of the network topology. This finding strongly suggests that defectors, despite their temporary advantages, are vulnerable to catastrophic cascading process. Cooperation becomes the optimal strategy to maximize benefit and to escape death. As a by-produce, our work resolved the social dilemma of profit versus cooperation in a natural manner.

Our results can yield insights into the mechanism of catastrophic events in economic and ecosystems. For example, a large scale bankruptcy of financial organizations may be a typical cascading process where high-risk investments, a kind of defection behavior, decrease the capacity of agents to resist deficiency and trigger the outbreak of the cascade. For evolutionary biology, our result may provide hints to the mechanism of massive species extinction in relatively short time scales.

Results were published in the following two papers:

1. W.-X. Wang, R. Yang, and Y.-C. Lai, "Cascade of elimination and emergence of pure cooperation in co-evolutionary games on networks," *Physical Review E (Rapid Communications)* **81**, 035102(R)(1-4)(2010).
2. W.-X. Wang, Y.-C. Lai, and D. Armbruster, "Cascading failures and the emergence of cooperation in evolutionary game based models of social and economical networks," *Chaos* **21**, 033112, 1-12 (2011).

3.3 Information explosion on complex networks and control

Spreading and transportation processes are fundamental and ubiquitous in a variety of complex systems: the Internet, biological networks, and social networks. Most previous works addressed how the underlying network structure affects the spreading and transportation dynamics, with efforts ranging from routing data traffic on the Internet to the spreading of epidemic opinions and rumors on either social networks or communication networks. Often, one focus of analysis and computation was on the asymptotic extent of the spreading process, as characterized by the percentage of the infected nodes after the termination of the process. In this regard, various processes such as those described by the two-state spreading model (SIS), the voter model, and the rumor-spreading model were studied. In most previous works, the entity of spreading, such as a particular type of virus or a piece of information, was assumed to be invariant during the process. In realistic situations, distortion of the entity during the spreading process can be expected, such as mutations of viruses, errors in transported data packets, and distorted opinion or rumors. The problem of information distortion is particularly relevant when human behaviors are involved in the spreading and transportation process. The distortions can lead to a significant increase, or even a divergence in the number of messages on the network over the time.

Information explosion has indeed occurred in the modern time. There has been an unprecedented growth in the number and variety of data collections as technology and network connectivity become increasingly affordable. Distortion in communication is inevitable and may contribute partially to the growth of data information. How to hold and release information becomes an issue of increasing importance, with implications ranging from personal privacy to national security.

We developed a model to address the problem of information explosion and control on complex networks. The starting points of our consideration were the following: (i) a node (or an agent) accepts or discards a message based on the existent information content in its memory, and (ii) information distortion can occur during the spreading process, which can be quantified with the probability p that a message is distorted after passing through an agent. In principle, the values of p can vary across different agents, but for simplicity we assumed that the spread in the probabilities is small and can be neglected. To gain insight, we examined the case where one message is set out to spread on the network initially. In the error-free case ($p = 0$), the number of messages is simply one. For p slightly above zero, the number of messages is greater than one. However, since p is small, a steady state can emerge where the average number of messages on the network tends to a constant. For large values of p, due to the frequent mutations, the number of distinct messages can increase with time. This introduces a positive feedback mechanism that generates an increasing amount of difficulty for agents to distinguish between the true and modified messages. As a result, different versions of the true message can accumulate in the memories of agents, generating even more distorted messages, and the number of messages can keep increasing with time, leading to information explosion. In general, as p is increased from zero and passes through a critical point p_c, a phase transition can occur from steady state to information explosion. Our main result was that this scenario can indeed occur on complex networks. Another result was with respect to a network's robustness to information explosion, which can be measured through the value of p_c, where a higher value indicates that the network is more robust. An issue was whether some control strategy can be derived to increase the network robustness. We demonstrated a process that controls an agent's selection of a neighbor to spread the message to, which can be used to maximize the value of p_c. All these were supported by a theoretical analysis of the controlled strategy of selection and extensive numerical computation.

The results were summarized in the following paper:

1. X.-J. Ma, W.-X. Wang, Y.-C. Lai, and Z.-G. Zheng, "Information explosion on complex networks and control," *European Journal of Physics B* **76**, 179-183 (2010).

3.4 Pattern formation, synchronization and outbreak of biodiversity in cyclically competing games

Biodiversity is ubiquitous in nature and fundamental to evolution in ecosystems. However, a significant challenge remains in understanding biodiversity since, by the principle of natural selection, only fitter species are supposed to survive from interactions and competitions with other species for limited resources. To resolve this dilemma, evolutionary game theory had been proposed as a paradigm to address the coexistence of competing species, which is the key to biodiversity.

A fundamental type of interactions in ecosystems is cyclic, non-hierarchical competitions. They have been observed in a plethora of real ecosystems ranging from microbes to mating strategies of side-blotched lizards in California. A paradigmatic system to study the role of the competitions in biodiversity is the classical, cyclic game of rock-paper-scissors. One approach is *macroscopic* in the sense that the mathematical models are aimed at describing the evolution of the populations of competing species, which are assumed to be well mixed. In this macroscopic approach, any species is treated as a whole through its population. An interesting result from this approach is that cyclic competitions alone are not sufficient to support species coexistence. The ubiquity of the coexistence phenomenon in nature suggests that additional factors must exist to promote coexistence and, consequently, biodiversity. To identify these additional factors and also to capture the complex interacting dynamics among individuals of competing species, *microscopic* game models incorporating stochastic interactions on spatially extended scales have been exploited with the remarkable result that, due to stochasticity and local interactions, coexistence can arise even in the presence of species dispersal. Since then, the role of mobility in coexistence in microscopic game models has been investigated, where it was found that strong *local* mobility can cause non-local interactions, which under certain circumstances tends to hamper coexistence through the formation of moving spiral waves of population densities in the physical space. The roles of epidemic spreading and intra-species competition in species coexistence, the basin structures, and competition in continuous physical space, were investigated by PI's group (papers #2 − 6 in the publication list below). An accepted notion in the field was that strong mobility is detrimental to biodiversity.

We uncovered a phenomenon that is in sharp contrast to this notion: species migration across vast spatial scales can in fact promote coexistence. Such movements are indeed common in ecosystems. To our knowledge, prior to our work, a microscopic understanding of the effect of large-scale migration on species coexistence had not been available. Since long-distance migrations can be regarded effectively as an extremely strong type of mobility, according to the conventional wisdom, coexistence would be disfavored or even prohibited. However, our studies revealed, strikingly, that migration favors coexistence and thereby promotes biodiversity.

We considered species movements on two distinct spatial scales: intra-patch and inter-patch migration, and studied microscopic stochastic games by focusing on the formation and the dynamics of self-organized patterns of species densities. Our microscopic model of inter-patch migration based on stochastic interactions was quite different from the coupled patchy models described by deterministic differential equations. We showed that the combination of intra- and inter-patch migrations can result in a robust type of coexistence characterized by the formation of a surprising class of *target* wave patterns, which had been found previously in different contexts (e.g., excitable systems). We found that, associated with coexistence, synchronization and time-lagged synchronization among spatial patterns in different patches can emerge, implying persistence of coexistence. An appealing feature of time-lagged synchronization is that it can potentially be used to anticipate the spatiotemporal evolution of species. We also found that the interplay between the two types of migration can result in a spontaneous outbreak of biodiversity in a world of single species with rare mutations. We established the robustness of the biodiversity-sustaining target waves with the aid of a basic concept in nonlinear dynamics: basins of attraction in the phase space. All the results were demonstrated using systematic simulations of microscopic game dynamics and substantiated by theoretical analysis based on nonlinear partial differential equations. Our results not only provided insights into the dynamics of global oscillations induced by long-distance interactions among cyclically competing species, but also had implications to the emergence of order from randomness and disorder in natural and social systems through self-organization in the absence of any central control.

The results were published in the following paper:

1. W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, "Pattern formation, synchronization and outbreak of biodiversity in cyclically competing games," *Physical Review E* **83**, 011917(1-9) (2011).

3.5 Predicting catastrophes in nonlinear dynamical systems by compressive sensing

Our basic idea to address the problem of predicting catastrophes in nonlinear dynamical systems can be described, as follows. We assume that an accurate model of the system is not available, i.e., the system equations are unknown, but the time evolutions of the key variables of the system can be accessed through monitoring or measurements. Our method consists of three steps: (i) predicting the dynamical system based on time series, (ii) identifying the parameters of the system, and (iii) performing bifurcation analysis using the predicted system equations to locate potential catastrophic events in the parameter space so as to determine the likelihood of system's drifting into a catastrophic regime. For example, if the system operates at a parameter setting close to such a critical bifurcation, catastrophe is imminent as a small parameter change or a random perturbation can push the system beyond the bifurcation point. Once a complete set of system equations has been predicted and the parameters have been identified, one needs to examine the available parameter space. In general, to explore the multi-parameter space of a dynamical system can be extremely challenging, which can often lead to the discovery of new phenomena in dynamics. The focus of our work, however, was on predicting the dynamical systems based on compressive sensing.

Our framework to fully reconstruct dynamical systems using time series alone was based on the assumption that the dynamics of many natural and man-made systems are determined by functions that can be approximated by series expansions in a suitable base. The major task is then to estimate the coefficients in the series representation. In general, the number of coefficients to be estimated can be large but many of them are zero (the sparsity condition). According to the conventional wisdom this would be a difficult problem as a large amount of data is required and the computations involved can be extremely demanding. However, the paradigm of compressive sensing developed in 2005-2006 provides a viable solution to the problem, where the key idea is to reconstruct a sparse signal from small amount of observations, as measured by linear projections of the original signal on a few predetermined vectors. Since the requirements for the observations can be considerably relaxed as compared with those associated with conventional signal reconstruction schemes, compressive sensing has become a powerful technique to obtain high-fidelity signal for applications where sufficient observations are not available. We articulated a general methodology to cast the problems of dynamical-system prediction into the framework of compressive sensing and we demonstrated the power of our method by carrying out bifurcation analyses on the predicted dynamical systems to locate potential catastrophes using exemplary chaotic systems.

Generally, the problem of compressive sensing can be described as the reconstruction of a sparse vector $\mathbf{a} \in R^v$ from linear measurements $\mathbf{X}$ about $\mathbf{a}$ in the form: $\mathbf{X} = \mathbf{G} \cdot \mathbf{a}$, where $\mathbf{X} \in R^w$, $\mathbf{G}$ is a $w \times v$ matrix and most components of $\mathbf{a}$ are zero. The compressive sensing theory ensures that the number of components of the unknown signal can be much larger than the number of required measurements for reconstruction, i.e., $v \gg w$. Accurate reconstruction can be achieved by solving the following convex optimization problem: $\min \|\mathbf{a}\|_1 \quad \text{subject to} \quad \mathbf{X} = \mathbf{G} \cdot \mathbf{a}$, where $\|\mathbf{a}\|_1 = \sum_{i=1}^{v} |\mathbf{a}_i|$ is the L_1 norm of $\mathbf{a}$.

We argued that the inverse problem of predicting dynamical systems can be cast in the framework of compressive sensing so that optimal solutions can be obtained even when the number of base coefficients to be estimated is large and/or the amount of available data is small. Here, we present a typical example to illustrate our method. Assume that the dynamical system can generally be written as $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, where $\mathbf{x} \in R^m$ represents the set of externally accessible dynamical variables and $\mathbf{F}$ is a smooth vector function in R^m. The jth component of $\mathbf{F}(\mathbf{x})$ can be represented as a power series:

$$[\mathbf{F}(\mathbf{x})]_j = \sum_{l_1=0}^{n} \sum_{l_2=0}^{n} \cdots \sum_{l_m=0}^{n} (a_j)_{l_1,\cdots,l_m} \cdot x_1^{l_1} x_2^{l_2} \cdots x_m^{l_m}, \tag{4}$$

where x_k ($k = 1, \cdots, m$) is the kth component of the dynamical variable, and the scalar coefficient of each product term $(a_j)_{l_1,\cdots,l_m} \in R$ is to be determined from measurements. Note that the terms in Eq. (4) are all possible products of different components with different powers, and there are $(1+n)^m$ terms in total.

It is useful to focus on one dynamical variable of the system. (Procedures for other variables are similar.) For example, to construct the measurement vector $\mathbf{X}$ and the matrix $\mathbf{G}$ for the case of $m = 3$ (dynamical variables x, y, and z) and $n = 3$, we have the following explicit dynamical equation for the first dynamical variable: $[\mathbf{F}(\mathbf{x})]_1 \equiv (a_1)_{0,0,0} x^0 y^0 z^0 + (a_1)_{1,0,0} x^1 y^0 z^0 + \cdots + (a_1)_{3,3,3} x^3 y^3 z^3$. We can denote the coefficients of $[\mathbf{F}(\mathbf{x})]_1$ by $\mathbf{a}_1 = [(a_1)_{0,0,0}, (a_1)_{1,0,0}, \cdots, (a_1)_{3,3,3}]^T$. Assuming that measurements of $\mathbf{x}(t)$ at a set of time $t_1, t_2, \ldots, t_w$ are available, we denote

$$\mathbf{g}(t) = \left[x(t)^0 y(t)^0 z(t)^0, x(t)^0 y(t)^0 z(t)^1, \cdots, x(t)^3 y(t)^3 z(t)^3\right],$$

such that $[\mathbf{F}(\mathbf{x}(t))]_1 = \mathbf{g}(t) \cdot \mathbf{a}_1$. From the expression of $[\mathbf{F}(\mathbf{x})]_1$, we can choose the measurement vector as $\mathbf{X} = [\dot{x}(t_1), \dot{x}(t_2), \cdots, \dot{x}(t_w)]^T$, which can be calculated from time series. Finally, we obtain the following equation in the form $\mathbf{X} = \mathbf{G} \cdot \mathbf{a}_1$:

$$\begin{pmatrix} \dot{x}(t_1) \\ \dot{x}(t_2) \\ \vdots \\ \dot{x}(t_w) \end{pmatrix} = \begin{pmatrix} \mathbf{g}(t_1) \\ \mathbf{g}(t_2) \\ \vdots \\ \mathbf{g}(t_w) \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \end{pmatrix}. \tag{5}$$

To ensure the restricted isometry property, we can normalize $\mathbf{G}$ through dividing elements in each column by the L_2 norm of that column: $(\mathbf{G}')_{ij} = (\mathbf{G})_{ij}/L_2(j)$ with $L_2(j) = \sqrt{\sum_{i=1}^{M}[(\mathbf{G})_{ij}]^2}$, so that $\mathbf{X} = \mathbf{G}' \cdot \mathbf{a}'_1$. After the normalization, $\mathbf{a}'_\mathbf{1} = \mathbf{a_1} \cdot L_2$ can be determined via some standard compressive-sensing algorithm. As a result, the coefficients $\mathbf{a_1}$ are given by $\mathbf{a}'_\mathbf{1}/L_2$. To determine the set of power-series coefficients corresponding to a different dynamical variable, say y, we can simply replace the measurement vector by $\mathbf{X} = [\dot{y}(t_1), \dot{y}(t_2), \cdots, \dot{y}(t_w)]^T$ and use the same matrix $\mathbf{G}$. This way all coefficients $\mathbf{a_1}$, $\mathbf{a_2}$, and $\mathbf{a_3}$ in three dimensions can be estimated.

We tested our prediction method using a number of physically relevant models of nonlinear dynamical systems in both discrete and continuous time, and demonstrated successful prediction of catastrophic bifurcations in all cases considered. A merit of our approach is that, due to the nature of the compressive-sensing method, a large number of terms can be accurately estimated even with short available time series, enabling potential implementation in real times. Predicting catastrophe is a problem of uttermost importance in science and engineering and of extremely broad interest as well, and our work represented a step forward in this area.

Details of this work can be found in

- W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, "Predicting catastrophes in nonlinear dynamical systems by compressive sensing," *Physical Review Letters* **106**, 154101, 1-4 (2011).

3.6 Time-series based prediction of complex oscillator networks via compressive sensing

Based on our work on predicting catastrophes nonlinear dynamical systems, we developed a framework that enables a full reconstruction of coupled oscillator networks whose vector field consists of a limited number of terms in some suitable base of expansion. The basic idea is that the mathematical functions determining the dynamical couplings in a physical network can be expressed as power-series expansions. The task is then to estimate all the nonzero coefficients. Since the underlying coupling functions are unknown, the power series can contain high-order terms. The number of coefficients to be estimated can therefore be quite large. However, the number of nonzero coefficients may be only a few so that the vector of coefficients is effectively sparse, rendering applicable compressive sensing.

Extensive computations revealed that both nonlinear nodal dynamics and node-to-node interactions can be accurately predicted, leading to reliable and robust reconstruction of the underlying networked system, as characterized by near-zero prediction errors, regardless of the nature of the nodal dynamics and the network structure. Although all the examples of nodal dynamics tested were polynomial vector fields, we examined other expansion bases such as trigonometric functions. If the prediction base is sufficiently wide to include all terms in the system equations as a small subset, high-accuracy prediction can be guaranteed, regardless of the mathematical forms of the terms in the equations. These features make our method appealing to predicting general complex networked systems with low data requirement.

This work was published as

- W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and M. A. F. Harrison, "Time-series based prediction of complex oscillator networks via compressive sensing," *Europhysics Letters* **94**, 48006, 1-6 (2011).

3.7 Reconstruction of social networks based on evolutionary-game data via compressive sensing

Evolutionary games are a common type interactions in a variety of complex, networked, natural and social systems. Given such a system, uncovering the interacting structure of the underlying network is key to understanding its collective dynamics. We articulated a general method to address the problem of uncovering network topology using evolutionary-game data based on compressive sensing. Although we had demonstrated that convex optimization, the essence of compressive sensing, can be used to construct coupled oscillator networks, we showed the advantage of compressive sensing, such as low data requirement, in solving the general inverse problem of network reconstruction based on either continuous or discrete data. In particular, in a typical game, agents use different strategies in order to gain the maximum payoff. The strategies can be divided into two types: cooperation and defection. We showed that, even when the available information about each agent's strategy and payoff is limited, our compressive-sensing based method can yield precise knowledge of the node-to-node interaction patterns in a highly efficient manner. We validated our method through (1) extensive numerical computations using model complex networks and evolutionary games, and (2) an actual social experiment in which participants forming a friendship network played a typical game to generate short sequences of strategy and payoff data. The high prediction accuracy achieved and the unique requirement of extremely small data set made our method appealing to potential applications to reveal "hidden" networks embedded in various social, economic and biological systems.

The mathematical formulation of our method is as follows. In an evolutionary game, at any time a player can choose one of two strategies S: cooperation (C) or defection (D), which can be expressed as $\mathbf{S}(C) = (1,0)^T$ and $\mathbf{S}(D) = (0,1)^T$. The payoffs of two players in a game is determined by their strategies and the payoff matrix of the specific game. For example, for the prisoner's dilemma game (PDG) and the snowdrift games (SG), the payoff matrices are

$$\mathbf{P}_{PDG} = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \text{ or } \mathbf{P}_{SG} = \begin{pmatrix} 1 & 1-r \\ 1+r & 0 \end{pmatrix}, \tag{6}$$

where b $(1 < b < 2)$ and r $(0 < r < 1)$ are parameters characterizing the temptation to defect. When a defector encounters a cooperator, the defector gains payoff b in the PDG and payoff $1+r$ in the SG, but the cooperator gains the sucker payoff 0 in the PDG and payoff $1-r$ in the SG. At each time step, all agents play the game with their neighbors and gain payoffs. For agent i, the payoff is

$$G_i = \sum_{j \in \Gamma_i} \mathbf{S}_i^T \mathbf{P} \mathbf{S}_j, \tag{7}$$

where $\mathbf{S}_i$ and $\mathbf{S}_j$ denote the strategies of agents i and j at the time and the sum is over the neighboring set Γ_i of i. After obtaining its payoff, an agent updates its strategy according to its own and neighbors' payoffs, attempting to maximize its payoff at the next round. Possible mathematical rules to capture an agent's decision making process include the best-take-over rule, the Fermi equation, and payoff-difference-determined updating probability. To be concrete, we used the Fermi rule in our simulations of evolutionary-game dynamics and generated time series accordingly, which is defined, as follows. After a player i randomly chooses a neighbor j, i adopts j's status $\mathbf{S}_j$ with the probability:

$$W(\mathbf{S}_i \leftarrow \mathbf{S}_j) = \frac{1}{1 + \exp\left[(G_i - G_j)/\kappa\right]}, \tag{8}$$

where κ characterizes the stochastic uncertainties in the game dynamics. For example, $\kappa = 0$ corresponds to absolute rationality where the probability is zero if $G_j < G_i$ and one if $G_i < G_j$, and $\kappa \to \infty$ corresponds to completely random decision. The probability W thus characterizes the bounded rationality of agents in society and natural selection based on relative fitness in evolution.

The key to solving the network-reconstruction problem lies in the relationship between agents' payoffs and strategies. The interactions among agents in the network can be characterized by an $N\times N$ adjacency matrix $\mathbf{A}$ with elements $a_{ij}=1$ if agents i and j are connected and $a_{ij}=0$ otherwise. The payoff of agent x can be written as

$$G_x(t)=a_{x1}\mathbf{S}_x^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_1(t)+\cdots+a_{x,x-1}\mathbf{S}_x^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_{x-1}(t)+a_{x,x+1}\mathbf{S}_x^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_{x+1}(t)+\cdots+a_{xN}\mathbf{S}_x^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_N(t), \tag{9}$$

where a_{xi} $(i=1,\cdots,x-1,x+1,\cdots,N)$ represents a possible connection between agent x and its neighbor i, $a_{xi}\mathbf{S}_x^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_i(t)$ $(i=1,\cdots,x-1,x+1,\cdots,N)$ stands for the possible payoff of agent x from playing game with i (if there is no connection between x and i, the payoff is zero because $a_{xi}=0$), and $t=1,\cdots,m$ is the number of round that all agents play the game with their neighbors. This relation provides us with a base to construct the measurement vector and the transform matrix in a proper compressive-sensing framework to obtain solution of the neighboring vector $\mathbf{A}_x$ of agent x. In particular, we can write

$$\begin{aligned}\mathbf{G}_x &= (G_x(t_1),G_x(t_2),\cdots,G_x(t_m))^T,\\ \mathbf{A}_x &= (a_{x1},\cdots,a_{x,x-1},a_{x,x+1},\cdots,a_{xN})^T,\end{aligned} \tag{10}$$

and $\sigma_x=$

$$\begin{pmatrix} F_{x1}(t_1) & \cdots & F_{x,x-1}(t_1) & F_{x,x+1}(t_1) & \cdots & F_{xN}(t_1)\\ F_{x1}(t_2) & \cdots & F_{x,x-1}(t_2) & F_{x,x+1}(t_2) & \cdots & F_{xN}(t_2)\\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots\\ F_{x1}(t_m) & \cdots & F_{x,x-1}(t_m) & F_{x,x+1}(t_m) & \cdots & F_{xN}(t_m)\end{pmatrix},$$

where $F_{xy}(t_i)=\mathbf{S}_x^T(t_i)\cdot\mathbf{P}\cdot\mathbf{S}_y(t_i)$. The vectors $\mathbf{G}_x$, $\mathbf{A}_x$ and the matrix σ_x satisfy

$$\mathbf{G}_x=\sigma_x\cdot\mathbf{A}_x, \tag{11}$$

where $\mathbf{A}_x$ is sparse due to the sparsity of the underlying complex network, making the compressive-sensing framework applicable. Since $\mathbf{S}_x^T(t_i)$ and $\mathbf{S}_y(t_i)$ in $F_{xy}(t_i)$ come from data and $\mathbf{P}$ is known, the vector $\mathbf{G}_x$ can be obtained directly while the matrix σ_x can be calculated from the strategy and payoff data. The vector $\mathbf{A}_x$ can thus be predicted based solely on the time series. Since the self-interaction term a_{xx} is not included in the vector $\mathbf{A}_x$ and the self-column $[F_{xx}(t_1),\cdots,F_{xx}(t_m)]^T$ is excluded from the matrix σ_x, the computation required for compressive sensing can be reduced. In a similar fashion, the neighboring vectors of all other agents can be predicted, yielding the network adjacency matrix $\mathbf{A}=(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_N)$.

We validated our method using model complex networks of different topologies. We also conducted an experiment to reconstruct a real friendship network. In the experiment, 22 participants from Arizona State University played PDG together iteratively and, at each round each player was allowed to change his/her strategies to optimize the payoff. The payoff parameter was set (arbitrarily) to be $b=1.2$. The player who had the highest normalized payoff (original payoff divided by the number of neighbors) summed over time was the winner and rewarded. During the experiment, each player was allowed to communicate only with his/her direct neighbors for strategy updating. Prior to experiment, there was a social tie (link) between two players if they had already been acquainted to each other; otherwise there was no link. Among the 22 players, two withdrew before the experiment was completed, so they were treated as isolated nodes. The network structure is illustrated in Fig. 1(a). It exhibits typical features of a social network, such as the appearance of dense triangles and a core consisting of 4 players (nodes 5, 11, 13, and 16), which is fully connected within and has more links than other nodes in the network. The core essentially consists of players who were responsible for recruiting other players to participate in the experiment. Each of the 20 players who completed the experiment played 31 rounds of games, and he/she recorded his/her own strategy and payoff

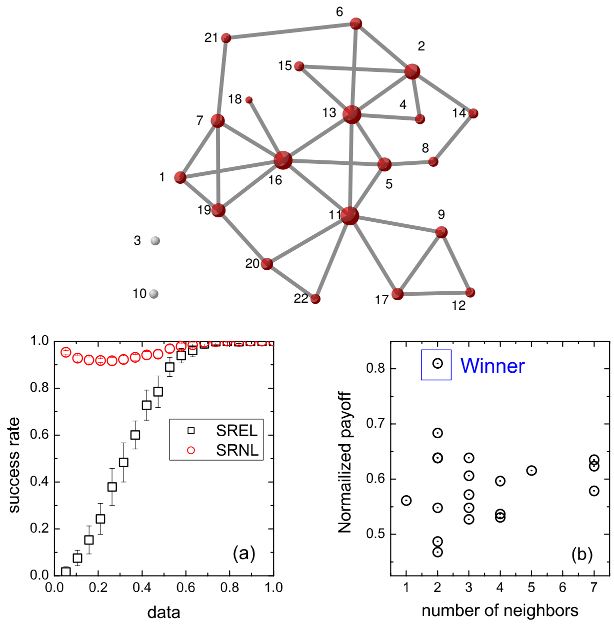

Figure 1: **Detection of a hidden node from a real social network**. (a) Structure of experimental social network. (b) Success rates of uncovering the network topology and (c) normalized payoff of player as a function of node degrees. The 10 independent realizations used in calculating the average success rates were randomly chosen from the data base of 31 rounds of games.

at each time, which represented the available data base for prediction. The data used for each prediction run was randomly picked from this data base. The pre-existed friendship ties among the participants tend to favor cooperation and preclude the system from being trapped in the social dilemma, due to the relatively short data streams. However, for a long run, a full defection state may occur. In this sense, the recorded data were taken during the transient dynamical phase and thus suitable for network reconstruction. The results are shown in Fig. 1(b). We see that the social network can be successfully uncovered, despite the complicated process of individual's decision making during the experiment.

An interesting phenomenon was that the winner picked in terms of the normalized payoff had only two neighbors, in contrast to the players with the largest node degree, whose normalized payoffs were approximately at the average level, as shown in Fig. 1(c). In addition, the payoffs of players of smaller degrees were highly non-uniform, while those of higher degrees showed smaller difference. This suggests that players of high degree may not act as leaders due to their average normalized payoffs. We also observed from experimental data that a typical player with a large number of neighbors failed to stimulate their neighbors to follow his/her strategies, suggesting that hubs may not be as influential in social networks. However, this finding should not be interpreted as a counter-example to the leader's role in evolutionary games, since the network based on friendship ties may violate the absolute selfish assumption of players who tend to be reciprocal with each other.

For all cases of networks hosting evolutionary-game dynamics that we considered, as the number of data points exceeds a low critical value depending on the sparsity of the underlying network, the prediction errors approach zero rapidly, without or with noise in the data. To our knowledge, no previous method could match our method in terms of the accuracy and efficiency, with only small set of discrete data. Our method, besides being fully applicable to complex networks governed by evolutionary-game type of interactions, can be applied to other contexts where the dynamical processes are discrete in time and the amount of available data is small. For example, inferring gene regulatory networks from sparse experimental data is a problem of paramount importance in systems biology. For such an application, Eq. (1) should be replaced by the Hill equation that models generic interactions among genes. In an expansion using base functions specifically suited for gene regulatory interactions, a compressive-sensing framework may be established. The underlying reverse-engineering problem can then be solved. A challenge that must be overcome is to represent the Hill function by an appropriate mathematical expansion so that the sparsity requirement for compressive sensing can be met.

The details of this work can be found in the following paper:

- W.-X. Wang, Y.-C. Lai, C. Grebogi, and J.-P. Ye, "Network reconstruction based on evolutionary-game data via compressive sensing," *Physical Review X* **1**, 021021, 1-7 (2011).

3.8 Optimizing controllability of complex networks by minimum structural perturbations

The ability to control complex networks is utter-mostly important to many critical problems in science, engineering and medicine, and has the potential to generate great technological breakthroughs as well. Indeed, because of the ubiquity of complex networks in natural, technological, social, and economical systems, it is highly desirable to be able to apply proper control to guide the network dynamics toward states with the best performance and, at the same time, to avoid undesired or deleterious states. While actual control of complex networks has not been achieved at the present, a necessary step is to understand the *controllability* of complex networks, which has become a topic of active pursuit. Specifically, given a complex-networked dynamical system, one wishes to assess whether it would be possible to apply certain number of control signals at an arbitrary set of nodes so as to drive the system toward some desirable state. The number of control signals, N_D, is thus a key quantity of interest as, qualitatively, it characterizes the cost to bring the system under control.

We started research on controllability of complex networks in 2011. The first question we asked was: given an arbitrary network that requires a certain number of signals to be controlled, can one perturb the network structure slightly so as to achieve the optimal controllability characterized by $N_D = 1$? The theoretical framework under which this question may be addressed is the minimum-input theory developed to characterize the controllability of networks with linear dynamics, which is based on classical control and graph theories. The basic goal of the minimum-input theory is to determine the minimum number of nodes to be driven externally to bring the whole network under control. According to this theory, only topological changes can alter the network controllability. To be illustrative, we investigated structural perturbation through adding links to the network to enhance its controllability. It is practically important to develop a paradigm that minimizes the number of added links to achieve $N_D = 1$; for otherwise optimal controllability can be achieved trivially by keeping adding links to the network until it becomes fully connected, which according to the minimum-input theory is fully controllable with a single input. Guided by this general consideration, we articulated a strategy to perturb the network by providing a minimum number of additional links at suitable locations determined by certain criterion. The performance of our perturbation scheme was compared with that in the case where links are randomly added to the network. Our optimization strategy bridged the network topology and controllability by providing useful insights into the effect of the former on the latter.

A detailed description of our structural perturbation strategy to optimize network controllability is as follows. According to Kalman's controllability rank condition, a canonical, linear, and time-invariant dynamical system, $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$, can be controlled from any initial state to any desired state in finite time, if and only if the $N \times NM$ controllability matrix $\mathbf{C}$ has full rank, i.e.,

$$\mathrm{rank}(\mathbf{C}) \equiv \mathrm{rank}\left[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \cdots, \mathbf{A}^{N-1}\mathbf{B}\right] = N \tag{12}$$

where $\mathbf{x} \in R^N$, $\mathbf{B}$ is the $N \times M$ input matrix, M is the number of driver nodes, and $\mathbf{u}(t)$ is a time-dependent input control vector. The full-rank condition (12) is appropriate for characterizing the controllability of network systems if A is the transpose of the adjacency matrix and N is the number of nodes. Of particular importance to our perturbation strategy is the concept of structural controllability, which can be used to

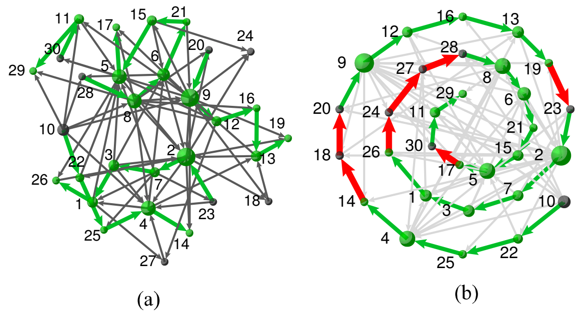

Figure 2: **Optimizing controllability of complex networks via structural perturbation**. (a) A network of 30 nodes with heterogeneous degree distribution, generated according to the preferential attachment mechanism. (b) All matching paths in order, starting from node 10 outside and ending at node 29 inside. The links of the set of maximum matching and the matched nodes are marked by green (gray). Structural perturbations are represented by the added links connecting the tail of a matching path in higher order to the head (black color) of a matching path in lower order, which are marked by red (dark gray). Other links are marked by light gray. The configuration of added links is not unique, but their minimum number is.

identify the minimum number N_D of driver nodes required for the system to satisfy the full-rank condition (12). However, it is practically difficult to check this condition for large complex networks, as the number of input combinations grows exponentially with the number of nodes ($\sim 2^N$). To overcome this difficulty, the Barabasi group proposed the concept of *maximum-matching set* to assess and quantify structural controllability. A key result is $N_D = 1$ if the network is perfectly matched; otherwise $N_D = N - N_M$, where N_M is the size of the maximum-matching set, i.e., the maximum set of links that do not share starting or ending nodes. The Barabasi group demonstrated that many real-world networks are far from being perfectly matched. Consequently, in order to fully control such a network, a large number of input signals applied to an equally large number of nodes are necessary. This motivated us to ask whether optimal control $N_D = 1$ is achievable by making deliberate, small structural perturbations to the network. We found that, given *any* network, a minimum number of links can indeed be added so that all nodes except one are matched. That is, under only one input control signal the perturbed network would meet the full-rank condition.

We formulated our strategy to optimize network controllability by adding minimum number of additional links for both directional and bidirectional networks. To explain our strategy, we introduced the concept of "matching path," a subset of links in the set of maximum matching (or "isolated" nodes), which can be (i) starting from an unmatched node and ending at a matched node without outgoing link belonging to the set of maximum matching, (ii) starting from an arbitrary node in a directed loop and ending at the "superior" node that points at the starting node, or (iii) an "isolated" node without any link belonging to the set of maximum matching. Here, case (ii) defines a "close matching path."

Our optimization process involves three steps: (1) finding the minimum number of independent matching paths, except close matching paths; (2) randomly ordering all found matching paths; (3) linking the ending points of each matching path to the starting nodes of the matching paths next to it in order, as illustrated in Fig. 2. The minimum number of independent matching paths, except close matching paths, is equal to one less than the number N_D of unmatched nodes. Applying such structural perturbations, the maximum fraction m_{max} of added links (m is the ratio of the number of added links to the number N_l of links in the original network) to achieve $N_D = 1$ is

$$m_{max} = \frac{N_D - 1}{N_l}. \tag{13}$$

If one external signal can control multiple drivers, the network will be fully controllable with a single controller imposed at the starting node of the first matching path and any one node in each of other close matching paths simultaneously. We proved, according to the classical structural controllability theory, that the optimal network resulted from the above structural perturbations satisfies the full-rank condition with a single input. The value of N_D can always be reduced to 1 by adding a minimum number of links.

In general, our method is applicable to networks for which establishing a link costs less than imposing a time-variant controller at a node, such as many technological and social networks. However, there are networks in the real world for which the opposite is true, such as gene regulatory networks, where to establish a new regulatory connection between genes may be more difficult than exogenously altering the expression of a gene. For such networks, our optimization method is not meaningful; alternative ways to enhance the network controllability must be explored. In addition, the issue of trade-off between network robustness in response to failures/attacks and lower control cost with less controllers can be critical.

The details can be found in the following paper:

- W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, "Optimizing controllability of complex networks by small structural perturbations," *Physical Review E* **85**, 026115, 1-5 (2012).

3.9 Detecting hidden nodes in complex networks from time series based on compressive sensing

When dealing with an unknown complex system that has a large number of interacting components organized hierarchically, curiosity demands that we ask the following question: are there hidden objects that are not accessible from the external world? The problem of inferring the existence of hidden objects from observations is quite challenging but it has significant applications in many disciplines of science and engineering. Here by "hidden" we mean that no direct observation of or information about the object is available, and so it appears to the outside world as a black box. However, due to the interactions between the hidden object and other observable components in the system, it may be possible to utilize "indirect" information to infer the existence of the hidden object and to locate its position with respect to objects that can be observed.

The paradigm of compressive sensing aims to reconstruct a sparse vector $\mathbf{a} \in \mathbb{R}^N$ from linear measurements $\mathbf{M}$ in the form $\mathbf{M} = \mathbf{G} \cdot \mathbf{a}$, where $\mathbf{M} \in \mathbb{R}^K$ and $\mathbf{G}$ is an $K \times N$ matrix. The compressive sensing theory guarantees that, when most components in the unknown vector $\mathbf{a}$ are zero, it can be reconstructed by fewer measurements than the number of components. The unknown vector $\mathbf{a}$ can be solved, for example, by a convex optimization procedure based on L_1 norm. Our work demonstrated that the problem of data-based network reconstruction can be casted into the form of $\mathbf{M} = \mathbf{G} \cdot \mathbf{a}$.

We considered networked systems for which the nodal dynamics, described by the vector function $\mathbf{F}_i(\mathbf{x}_i)$, can be separated from the interactions or coupling with other nodes in the network, mathematically described by the coupling function $\mathbf{H}_{ij}(\mathbf{x}_i, \mathbf{x}_j)$. The system can then be written as $\mathbf{M}_i = \mathbf{F}_i(\mathbf{x}_i) + \sum_{j \neq i}^{N} w_{ij} \mathbf{H}_{ij}(\mathbf{x}_i, \mathbf{x}_j)$, where $\mathbf{M}_i$ is the system response, either in discrete or continuous time. For example, for discrete-time mapping system, $\mathbf{M}_i$ are the state variables at the next time step, while in continuous system $\mathbf{M}_i$ are the derivatives of the corresponding variables. To illustrate our method to detect hidden nodes in a concrete manner, we assumed that the nodal and coupling functions can be written as some series expansion, e.g., power or Fourier series. In particular, we wrote: $\mathbf{F}_i(\mathbf{x}_i) = \sum_\gamma \tilde{a}_i^{(\gamma)} \tilde{g}_i^{(\gamma)}(\mathbf{x}_i)$ and $\mathbf{H}_{ij}(\mathbf{x}_i, \mathbf{x}_j) = \sum_\beta a_{ij}^{(\beta)} g_{ij}^{(\beta)}(\mathbf{x}_i, \mathbf{x}_j)$, where $\tilde{g}_i^{(\gamma)}$ are the expansion bases associated with $\mathbf{x}_i$ only, and $g_{ij}^{(\beta)}$ are with respect to both $\mathbf{x}_i$ and $\mathbf{x}_j$. Next we combined the bases $\tilde{\mathbf{g}}_i(t)$ and $\mathbf{g}_{ij}(t)$ at time t into a row vector, and the coefficients $\mathbf{a}_i^{(\alpha)}$ and $\mathbf{a}_{ij}^{(\beta)}$ into a constant column vector. The time-series vector of responses $\mathbf{M}_i(t)$ for node i can then be expressed by the

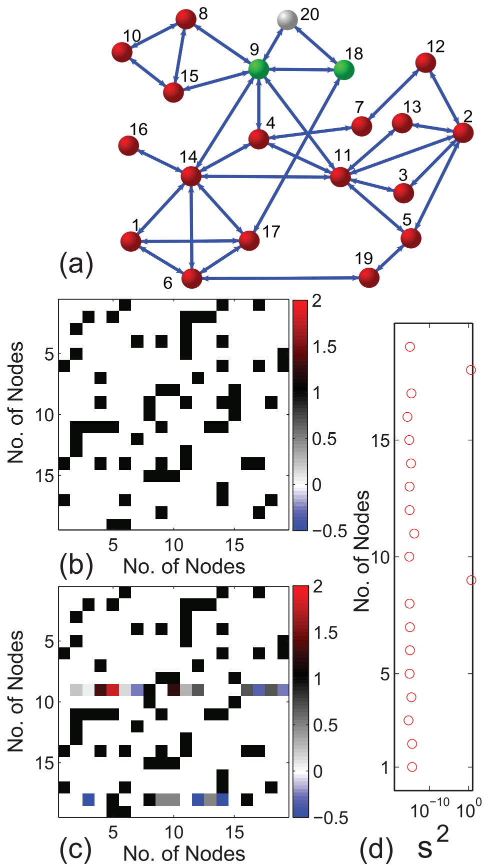

Figure 3: **Detection of a hidden node**. (a) Illustration of a complex network with a hidden node. (b) Representation of the true adjacency matrix, (c) reconstructed adjacency matrix elements for nodes except the hidden node based on time series from these nodes. (d) Variance σ^2 of the reconstructed coefficient vector **a** for all nodes, calculated by using 10 different random segments from the available experimental time series. The variances of the two green nodes (No. 9 and No. 18) are much larger than those of the red nodes, indicating that they are the neighbors of the hidden node.

product of the matrix $\mathbf{G}_i$ and the *to-be-determined* coefficient vector $\mathbf{a}_i$, with $\mathbf{G}_i$ given by

$$\mathbf{G}_i = \begin{pmatrix} \tilde{\mathbf{g}}_i(t_1) & \mathbf{g}_{i1}(t_1) & \cdots & \mathbf{g}_{ij}(t_1) & \cdots & \mathbf{g}_{iN}(t_1) \\ \tilde{\mathbf{g}}_i(t_2) & \mathbf{g}_{i1}(t_2) & \cdots & \mathbf{g}_{ij}(t_2) & \cdots & \mathbf{g}_{iN}(t_2) \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \tilde{\mathbf{g}}_i(t_m) & \mathbf{g}_{i1}(t_m) & \cdots & \mathbf{g}_{ij}(t_m) & \cdots & \mathbf{g}_{iN}(t_m) \end{pmatrix}, \tag{14}$$

where $\tilde{\mathbf{g}}_i(t)$ is the set of bases of $\mathbf{F}_i(\mathbf{x}_i)$, and $\mathbf{g}_{ij}(t)$ is the set of expansion bases of $\mathbf{H}_{ij}(\mathbf{x}_i, \mathbf{x}_j)$. Elements in the vector $\mathbf{M}_i(t)$ contain system response $m_i(t)$ at different t. In particular, when the vector $\mathbf{a}_i$ is determined via solving $\mathbf{M} = \mathbf{G} \cdot \mathbf{a}$, the dynamical equations for the set of corresponding variables at all nodes become known. Note that the vector $\mathbf{a}_i$ contains all the coupling weights from other nodes to i as in $\mathbf{g}_{ij}(t)$ and the complete information about the nodal dynamical equations as in $\tilde{\mathbf{g}}_i(t)$. Our earlier works had demonstrated that solutions to the compressive sensing problem can be obtained but only when time series from *all* nodes are available, i.e., when there is no hidden object.

To devise a compressive-sensing based methodology for detecting hidden nodes, we considered the case of one hidden node (or one cluster of hidden nodes). Let node i be one of the immediate neighbors of the hidden node. Due to lack of time series from the hidden node, the form $\mathbf{M} = \mathbf{G} \cdot \mathbf{a}$ is violated for node i, despite the available time series from other nodes in the network. That is, due to the missing time series from the hidden node and consequently missing elements in **a**, it is not possible to obtain the true solution of the dynamical equations of node i. If a node does not neighbor any hidden node, time series from itself and all its direct neighbors are available, rendering valid the form $\mathbf{M} = \mathbf{G} \cdot \mathbf{a}$ for such a node. The practical importance is that the errors in the prediction of the dynamics of the immediate neighbors of the hidden node will be much larger than those associated with nodes that do not have any hidden node in their neighborhoods. The predicted characteristics of all neighboring nodes of the hidden node will then show significant anomalies as compared with those of other nodes. The anomalies can be used to identify all nearest neighbors of the hidden node, which in turn imply its existence and its position in the network.

While our general idea of detecting hidden nodes can be formulated using different types of dynamical systems, to be concrete we describe here how this can be done using evolutionary-game type of dynamics. Such dynamical processes can be used to model generic agent-to-agent interactions in economical, social, or even certain biological networks. In an evolutionary-game system, the neighbors of the hidden node can be identified by utilizing the stability criterion with respect to different measurements. More specifically, in an evolutionary-game system, at any time a player can take on one of two strategies: cooperation (C) or defection (D), mathematically represented as $\mathbf{S}(C) = (1,0)^T$ and $\mathbf{S}(D) = (0,1)^T$, respectively. The payoffs of the two players in a game are determined by their strategies and the payoff matrix $\mathbf{P}$. For example, for the classical prisoner's dilemma game (PDG), the matrix elements are $P_{11} = 0$, $P_{12} = 0$, $P_{21} = b$, and $P_{22} = 0$, where $1 < b < 2$ is a parameter characterizing the temptation to defect. At each time step, all agents in the network play the game with their neighbors simultaneously and gain rewards. For agent i, the reward is $m_i = \sum_j a_{ij}\mathbf{S}_i^T\mathbf{P}\mathbf{S}_j$, where $\mathbf{S}_i$ and $\mathbf{S}_j$ denote the strategies of agents i and j taken at the time and a_{ij} is the coupling strength between them. After obtaining its payoff, an agent updates its strategy according to its own and neighbors' payoffs, attempting to maximize its payoff at the next round. We assumed that the strategy and payoff data of agents are available except those of the hidden node. In particular, we chose $\mathbf{g}_{ij}(t) = \mathbf{S}_i^T(t)\cdot\mathbf{P}\cdot\mathbf{S}_j(t)$ and ignore $\tilde{\mathbf{g}}_i$. The payoff of node i at different time t can be expressed as $\mathbf{M}_i(t) = \mathbf{G}_i\cdot\mathbf{a}_i$, where $\mathbf{G}_i$ is to be constructed as specified in Eq. (14), and the vector $\mathbf{a}_i$ to be determined contains all interaction strength between nodes i and other accessible nodes in the network. The network structure is uncovered after $\mathbf{a}$'s for all nodes are determined.

As an example, we obtained results of experimentally detecting a hidden node from a social network hosting evolutionary-game dynamics. In the experiment, 20 participants from Arizona State University played the PDG iteratively, with a pre-specified payoff parameter. The player with the highest normalized payoff (total payoffs normalized by their degrees) summed over time was the winner. The players can gamble with all their nearest neighbors in the pre-existing social network [Fig. 3(a)]. The network was determined by surveying the friendships among those participants, and it exhibits some typical properties of real social network, such as the much larger degree in some hub nodes. During the experiment, the strategies of each player and the gained payoff were recorded in all 32 rounds, except for the hidden node No. 20. The true adjacency matrix of the accessible nodes is represented in Fig. 3(b), and the predicted matrix is shown in Fig. 3(c). We see that the links of the two neighboring nodes (No. 9 and No. 18) of the hidden node No. 20 cannot be reliably predicted. Especially, the two nodes are predicted to have links with almost all nodes in the network, which is highly unlikely for a random network that is typically sparse. While the predicted loss of sparsity of certain nodes is an indication that they might be in the neighborhood of some hidden node, the condition is not sufficient in general, because of the existence of hub nodes with significantly more links than average in a complex network. Other conditions must then be sought in order to identify the neighbors of the hidden nodes. Our idea was to exploit the stability of the predicted solution with respect to different measurements used for compressive sensing. In particular, for the neighboring nodes of the hidden node, due to the lack of information needed to solve the underlying compressive-sensing problem, when different segments of the time series are used, the algorithm will yield different coefficient vectors $\mathbf{a}$. However, for a node not in the immediate neighborhood of the hidden node, the predicted vector $\mathbf{a}$ should be the same for different data segments, since the corresponding coefficients with the hidden node are zero. As shown in Fig. 3(d), the variances in $\mathbf{a}$ of nodes No. 9 and No. 18 from a number of predictions are much larger than those (essentially zero) of other nodes. Violation of sparsity in combination with the instability of the predicted solution then allows us to identify all neighbors of the hidden node, and consequently itself, with high confidence.

Details of this work can be found in

- R.-Q. Su, W.-X. Wang, and Y.-C. Lai, "Detecting hidden nodes in complex networks from time series," *Physical Review E (Rapid Communication)* **85**, 065201(R), 1-4 (2012).

3.10 Forecasting synchronizability of complex networks from data

The most amazing feature of a complex dynamical system consisting of a large number of interacting units (or components) is the emergence of collective dynamics. Indeed, it is this feature of "more is different" which makes complex systems extremely interesting and the study of collective dynamics fundamentally important to many natural and technological systems. Given a complex system, if the underlying mathematical rules or equations are completely known, then *in principle* the possible types of collective dynamics in the system can be predicted and studied, and most existing works on complex systems are of this nature. In realistic applications one may encounter the situation where, for a complex system of interest, the local system equations and the interactions among the components are not known *a priori* but only a set of time series are available. Can one still forecast or anticipate whether a certain type of collective dynamics can potentially occur in the system?

Even when the system equations of a complex system are known, it is still extremely challenging to predict, investigate, and exploit the emergence and evolution of collective dynamics. In order to address the issue of time-series based prediction of collective dynamics, it is useful to focus on a relatively well known class of such dynamics. Specifically, we studied coupled-oscillator networks, a paradigm for probing and understanding the synchronous behavior of interacting units with nonlinear dynamics. When the system equations are known, a widely used tool to determine whether synchronization can emerge physically is the master-stability function (MSF). In the MSF framework, synchronization under various combinations of network structures and oscillator dynamics can be predicted. For example, given the nodal dynamical equations, possible states of synchronization can be determined, which are basically the possible dynamics on the synchronization manifold. The MSF is nothing but the largest Lyapunov exponent characterizing the transverse stability of the synchronous dynamical state. For a typical nonlinear or chaotic oscillator, our previous research revealed that there may exist an open interval in the space of some generalized coupling parameter, where the MSF is negative so that any point in this interval can lead to stable synchronization. When the network structure is given, the set of eigenvalues of the underlying coupling matrix can be determined. For a network of coupled oscillators, the phase-space dimension can be extremely high, so there can be many transverse subspaces. The set of eigenvalues, after suitable normalization, gives the set of effective generalized coupling parameters associated with all the transverse subspaces. Network synchronization can occur only when all these parameters fall into the interval of negative MSF.

We proposed a general approach to forecasting the emergence of synchronization in complex oscillator networks based on time series. The specific setting of the problem is, as follows. Assume that at the time of interest the oscillator network is in an asynchronous state and time series from each node in the network can be obtained. Assume further that there exists a parameter characterizing the average coupling strength among the nodes. The question we asked was whether it would be possible to predict that synchronization can or cannot occur when the coupling parameter is allowed to change. Our method consists of two steps. First, we reconstructed the full topology of the network, together with the coupling strengths and the nodal dynamics, based solely on time series. This is accomplished by casting the prediction or reverse-engineering problem into the framework of compressive sensing. Here the relevant vector to be reconstructed originated from both nodal dynamics and topology, which is typically sparse due to the sparsity of complex networks. Second, from the predicted nodal dynamics and network structure, we performed synchronizability analysis by using the standard MSF approach. We validated our method by using random *weighted* networks of both continuous-time and discrete-time chaotic systems (e.g., the classical Lorenz system and Hénon map. Our computation and analysis indicated that with only small amount of measured data, the synchronization regions in the parameter space as identified by MSF and the network structure can be accurately predicted, rendering possible inference of synchronous dynamics. The critical data requirement and sampling frequency for different network sizes and degree distributions were studied in detail. The issue of the effect of measurement noise on prediction accuracy was addressed. In addition, the dependence of data requirement and computational time on the network size were studied.

One potential application of our prediction method is to control coupled oscillators to bring the system to synchronization. The base of control is prediction of future behavior by decoding the available time series at the present. If the natural dynamics in the future are undesirable, one can implement certain control scheme to drive the system to avoid the undesirable state before it emerges. This, however, requires relatively complete knowledge about the networked dynamical system, which can be achieved by exploiting the compressive-sensing paradigm. Consider the situation where synchronization is a desirable state of operation for the system, assuming that the system is not synchronized at the present. The first step is to determine, from currently available time series, whether synchronization is intrinsically likely to emerge. An answer can be obtained by using the reconstructed network structure and dynamics to estimate the network eigenvalue spectrum and MSF. The answer can be affirmative, for example, if the MSF is predicted to be negative in an open generalized coupling-parameter interval. That the system is not currently synchronized indicates that the normalized eigenvalue spectrum does not fall into the interval and, hence, suitable control can be applied to rescale and shift the eigenvalue spectrum into the negative MSF interval. To illustrate this method, we used the network system of coupled chaotic Lorenz oscillators. Figure 4(a) shows some representative time series in a case where the network is not synchronized, and the corresponding MSF and eigenvalue spectrum calculated from the reconstructed network structure and dynamics are shown in Fig. 4(c). It can be seen that some values of coupling parameter K [data points in Fig. 4(c)], the product between the coupling strength ξ and eigenvalues μ, are not located in the synchronizable region as indicated by the MSF [curve in Fig. 4(c)]. Thus, at the current parameter setting, synchronization cannot be realized in the system. In order for synchronization to emerge, all K values must fall into a region where the MSF is negative. A simple and practical way to manipulate K is to adjust the coupling strength but to keep the nodal dynamics and network structure unchanged. When the coupling strength ξ is modified, the network system can indeed achieve synchronization, as shown by the synchronous time times in Fig. 4(b). Examination of the MSF and eigenvalue spectrum indicated that, indeed, in this case all K values fall into the negative MSF interval. It should be emphasized that a prerequisite to this simple control scheme is full knowledge of the network structure and dynamics which, as we demonstrated, can be faithfully reconstructed based solely on small amount of data.

Details of this work can be found in

- R.-Q. Su, X. Ni, W.-X. Wang, and Y.-C. Lai, "Forecasting synchronizability of complex networks from data," *Physical Review E* **85**, 056220, 1-11 (2012).

3.11 Emergence of grouping in multi-resource minority game dynamics

The Minority Game (MG) was originated from the El Farol bar problem in game theory first conceived by Arthur in 1994, where a finite population of people try to decide, at the same time, whether to go to the bar on a particular night. Since the capacity of the bar is limited, it can only accommodate a small fraction of all who are interested. If many people choose to go to the bar, it will be crowded, depriving the people of the fun and thereby defying the purpose of going to the bar. In this case, those who choose to stay home are the winners. However, if many people decide to stay at home then the bar will be empty, so those who choose to go to the bar will have fun and they are the winners. Apparently, no matter what method each person uses

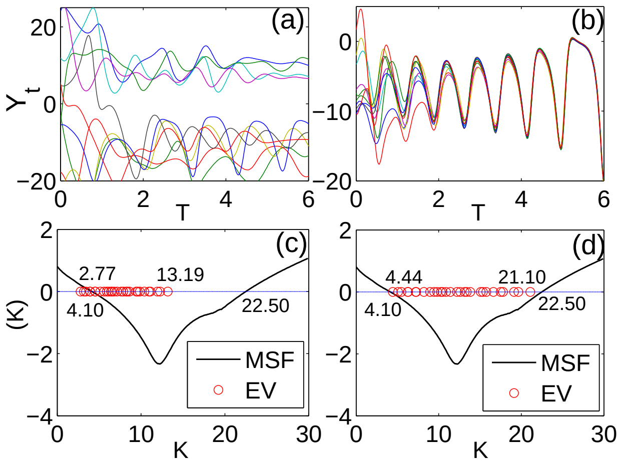

Figure 4: **Optimizing synchronizability of complex networks via prediction and control**. (a,b) Time series from 10 of the $N = 30$ nodes in two random networks of global coupling strength $\xi = 1$ and $\xi = 1.6$, respectively. The network is not synchronized in (a) but there is synchronization in (b). Other parameters are the same for both bases: connection probability $p = 0.2$ and the weight distribution interval is $[0.9, 1.0]$. (c,d) Rescaled eigenvalues $K_i (= \xi\mu_i)$ (denoted by open circles) of the network coupling matrices with respect to the MSF (denoted by solid lines) inferred from the same nodal dynamics and coupling scheme from the time series in (a,b), respectively.

to make a decision, the option taken by majority of people is guaranteed to fail and the winners are those that choose the minority strategy. Indeed, it can be proved that, for the El Farol bar problem there are mixed strategies and a Nash-equilibrium solution does exist in which the option taken by minority wins. A variant of the problem was subsequently proposed by Challet and Zhang, where a player among an odd number of players chooses one of the two options at each time step. Subsequently, the model was studied in a series of works. In physics, MG has received a great deal of attention from the statistical-mechanics community, especially in terms of problems associated with non-equilibrium phase transitions.

In the literature, the setting of MG is that there is a *single* resource but players have two possible strategies (e.g., in the El Farol bar problem there is a single bar and the two strategies are going to the bar or staying at home), and an agent is assumed to react to available global information about the history of the system by taking on an alternative strategy that is different than its current one. An outstanding problem concerned about the nonlinear dynamics of MG with *multiple resources*. We developed a class of multi-resource MG models. In particular, we assumed that, at any time, an individual agent has $k > 1$ *resources/strategies* to choose from. We introduced a parameter p, which is the probability that each agent reacts based on the available local information by selecting a less crowded resource in an attempt to gain higher payoff. We assumed realistically that only local information about the immediately preceding step is available, which constitutes the input to the model. This differs from the original MG model where global information is assumed to be available to all the agents and they make actions based on the past history. Letting p be the *minority-preference probability*, we found that, as p is increased, the striking phenomenon of grouping emerges, where the resources can be distinctly divided into two groups according to the number of their attendees. In addition, the number of stable pairs of groups also increases. We demonstrated the phenomenon numerically and derived an analytic theory to fully explain the phenomenon. We also showed that the grouping phenomenon plays a fundamental role in shaping the fluctuations of the system. An application to a real financial-market system by analyzing the available empirical data was also demonstrated, where grouping of stocks (resources) emerges. Our model is not only directly relevant to nonlinear and complex dynamical systems, but also applicable to social and economical systems.

The bifurcation-like phenomena associated with resource grouping in minority-game systems are not limited to the double-grouping (or paired grouping) behavior. In fact, we also observed phenomena such as period-3 double-grouping and period-3 triplet-grouping bifurcation. Further efforts are justified to explore various nonlinear dynamical phenomena in minority-game type of systems that describe a large variety of social, economical, and political systems.

Details of this work can be found in

- Z.-G. Huang, J.-Q. Zhang, J.-W. Dong, L. Huang, and Y.-C. Lai, "Emergence of grouping in multi-resource minority game dynamics," *Nature Scientific Reports* **2**, 703, 1-8 (2012).

3.12 Optimizing cooperation on complex networks in the presence of failure

Natural selection favors the survival and prevalence of species with competitive edge, yet the phenomenon of cooperation is ubiquitous in many biological, economical, and social systems. Understanding the emergence and evolution of cooperation has thus become a field of significant interdisciplinary interest, where evolutionary-game theory has served as a powerful mathematical paradigm. In a typical setting, a number of agents on a network interact with one another, where the network topology can be regular or complex, and each agent can take on one of the two strategies at any given time: cooperation or defection. The defection strategy is a selfish action that usually generates higher payoff temporally, as in paradigmatic games such as the PDGs, the snowdrift games (SGs), and the public goods games (PGGs). A basic issue is then how cooperation can possibly survive when natural selection favors the defection strategy in order to gain higher individual fitness (at least temporally). In the past two decades, many cooperation-facilitating mechanisms were uncovered, which include network reciprocity, reputation and punishment, random diffusion, success-driven migration, memory effect, benefit of noise, social diversity, asymmetric cost, and teaching ability.

In most previous works, no death mechanism was incorporated in the evolutionary-game model on networks, i.e., no agent can be removed from the system, even if it gains no payoff in a substantial amount of time. In real-world situations, an agent can go bankrupt and be eliminated immediately when its payoff falls below a critical threshold for certain period of time. An example is the great economical recession in 2008, where a large number of financial institutions and corporations collapsed. In an ecological system, death of individuals is a common phenomenon. In this regard, our previous work incorporated a simple elimination mechanism into the evolutionary-game rules. In particular, a tolerance parameter was assigned to each individual in the network, which is the lowest allowed payoff. An agent dies and is removed from the network when its payoff falls below this threshold. The threshold can be heterogeneously distributed among agents. It had been shown that rapid, cascading-like elimination of agents can result from such a death mechanism, and a pure cooperation state can emerge afterwards, where all defectors are eliminated and the survivors are exclusively cooperators. One implication is that defectors, despite their advantages in getting temporarily higher payoffs, may be particularly vulnerable to large-scale, catastrophic failures. These findings thus suggest that, in a complex system where agents are subject to failure or death, cooperation may be beneficial to mitigating large-scale breakdown.

During the performance period, we developed a control scheme to enhance cooperation and eliminate large-scale failures in complex networked systems. Our key idea was that, due to the complex time evolution of the system, although the payoff of any agent can inevitably become arbitrarily low, the probability that the payoff remains low for an extended period of time will be small. We were thus led to introduce a *time tolerance* for each agent, where an agent will not die or be removed unless its payoff remains below a critical threshold for time longer than the tolerance. Since the degree distribution of the network is in general not uniform, it is reasonable that the time tolerance be degree-dependent. A parameter β can then be introduced to characterize the heterogeneity of the distribution of the time tolerance, where $\beta = 0$ signifies completely uniform distribution. Our main result was that properly chosen time delay can optimize cooperation and prevent large-scale death. A surprising finding was that optimal state of cooperation occurs near $\beta = 0$, indicating that making time-tolerance distribution uniform is an effective strategy to enhance cooperation.

Specifically, to impose a time tolerance on a complex network, we conceived that a node or an agent's *debt capacity* depends on its relative "importance" in the network. We thus hypothesized the following relationship between agent i's time tolerance and its degree k_i:

$$T_i = NT_0 \frac{k_i^\beta}{\Sigma_l k_l^\beta}, \tag{15}$$

where N is the total number of agents, T_0 is the nominal time tolerance, and β is an externally control parameter. For $\beta < 0$, agents with higher (lower) degree have lower (higher) time tolerance, the situation is the opposite for $\beta > 0$, and $\beta = 0$ corresponds to uniform time tolerance in the network. A large values of T_i means that the node is more resilient to failure or "bankruptcy." A death mechanism can be introduced, e.g., for PDG by choosing the following payoff tolerance for agent i $P_i^T \equiv \alpha P_i^N = \alpha k_i$, where agent i dies and is removed from the network if its payoff is lower than P_i for consecutive T_i time steps, P_i^N is the normal payoff of agent i when the system is in a healthy state in which all agents are cooperators, and $0 < \alpha < 1$ is a tolerance parameter. Since an agent's degree may change when their neighbors die, k_i is the "instantaneous" degree of agent i. For $\alpha = 1$, agents have zero payoff tolerance to breakdown, while for $\alpha = 0$, agents are completely tolerant.

In our evolutionary game model, each time step (iteration) thus consists of the following four dynamical processes. (1) *Game playing and payoffs.* Each agent plays the classical PDG with all its nearest neighbors, and the total payoff is the sum of the payoffs gained in its two-player games with all other connected agents. The PDG parameters are chosen to be $R = 1$, $T = b > 1$, and $S = P = 0$. (2) *Strategy updating.* At each time step, agent i randomly chooses a neighbor j and imitates j's strategy with the probability $W_{i\to j} = \{1 + \exp[-(P_j - P_i)/\kappa]\}^{-1}$, where P_i and P_j are the payoffs agents i and j, and κ is the level of agents' "rationality" representing the uncertainties in assessing the best strategy. We set $\kappa = 0.1$. (3) *Failure and agent removal.* At each iteration, for agent i, the time in debt t_i increases by 1, if P_i falls below the payoff tolerance P_i^T during the prior t_i time steps. Otherwise, we set $t_i = 0$. Since k_i varies with time, P_i^T and T_i also change with time. If $t_i > T_i$ or if $k_i = 0$, agent i and all its links will be removed from the network. (4) *Random rewiring.* For agent i whose neighbor j has been removed in step (3), a new connection is added between agent i and an randomly selected agent in the remaining agents outside i's current neighborhood, provided that such an agent exists. This is motivated by the consideration that an agent in general will try to seek and engage new partners when the payoffs of some of its current partners become insignificant, and lack of global information leads to random selection. Note that, dynamical processes (1) and (2) are conventional for typical evolutionary-game dynamics, but processes (3) and (4) are unique features of our model.

Our computation and heuristic analysis indicated that, despite the network's being highly heterogeneous, making the time tolerance as uniformly as possible across the network can lead to the emergence of a stable cooperation cluster that has recruited majority of the agents in the network. Simultaneously, substantial death of agents can be avoided. This finding may have implications to policy making to prevent, for example, large-scale breakdown of social and economical systems. The emergence and evolution of cooperation in complex systems have been recognized as a fundamental issue in natural, social, and economical sciences, and our work may provide insights into the control of complex dynamical systems in terms of critical issues such as stability, performance, and sustainability.

The details can be found in

- Y.-Z. Chen and Y.-C. Lai, "Optimizing cooperation on complex networks in the presence of failure," *Physical Review E (Rapid Communications)* **86**, 045101(R), 1-4 (2012).

3.13 Exact controllability of complex networks

One of the most challenging problems in modern network science and engineering is controlling complex networks. While great effort had been devoted to understanding the interplay between complex networks and dynamical processes taking place on them in various natural and technological systems, control of complex dynamical networks remained to be an outstanding problem. Generally, because of the ubiquity of nonlinearity in nature, one must consider control of complex networked systems with nonlinear dynamics. However, at present there is no general framework to address this problem because of the extremely complicated interplay between network topology and nonlinear dynamical processes, despite the development of nonlinear control in certain particular situations such as consensus, communication, traffic and device networks. To ultimately develop a framework to control complex and nonlinear networks, a necessary and fundamental step is to investigate the controllability of complex networks with linear dynamics. There existed well developed theoretical frameworks of controllability for linear dynamical systems in the traditional field of engineering control. However, significant challenges arise when applying the traditional controllability framework to complex networks due to the difficulty to determine the minimum number of controllers. A ground-breaking contribution was made by the Barabasi group in 2011 who developed a minimum input theory to efficiently characterize the structural controllability of complex networks, allowing a minimum set of driver nodes to be identified to achieve full control. In particular, the structural controllability of a directed network can be mapped into the problem of maximum matching, where external control is necessary for every unmatched node. The structural-controllability framework also allows several basic issues to be addressed, such as linear edge dynamics, lower and upper bounds of energy required for control, control centrality, and optimization.

Although the structural-controllability theory offers a general tool for controlling directed networks, a universal framework for addressing the controllability of complex networks with arbitrary structures and configurations of link weights had been missing. Mathematically, the framework of structural controllability is applicable to directed networks characterized by structural matrices, in which all links are represented by independent free parameters. This requirement may be violated if exact link weights are given, motivating us to pursue an alternative framework beyond the structural-controllability theory. For undirected networks, the symmetric characteristic of the network matrix accounts for the weak violation of the assumption of structural matrix, even with random weights. Thus we continued to lack a reliable tool to measure the controllability of undirected networks. For some practical issue towards achieving actual control, such as predicting control energy given link weights, necessary and sufficient conditions to ensure full control are prerequisite. Taken together, a more general and accurate framework to study the controllability of complex networks was needed.

Supported by AFOSR, we developed an exact-controllability framework as an alternative to the structural-controllability framework, which offers a universal tool to treat the controllability of complex networks with arbitrary structures and link weights, including directed, undirected, weighted and unweighted networks with or without self-loops. Structural controllability can be reproduced in our framework for structural matrix that can be ensured by assigning random weights to directed links. In particular, based on the Popov-Belevitch-Hautus (PBH) rank condition that is equivalent to the Kalman rank condition, we proved that the minimum number of independent driver nodes or external controllers is equal to the maximum geometric multiplicity of all eigenvalues of the network matrix. If the network matrix is diagonalizable, e.g., as for undirected networks, controllability is simply determined by the maximum algebraic multiplicity of all eigenvalues. That is, the minimum number of inputs is determined by the dimension of eigenvectors for arbitrary networks and, for symmetric networks, this number is nothing but the eigenvalue degeneracy. For simple regular networks, their exact controllability can be calculated analytically. For more complicated model networks and many real-world weighted networks with distinct node-degree distributions, the exact controllability can be efficiently assessed by numerical computations. The minimum set of driver nodes can be identified by elementary transformation based on the exact-controllability framework. Our systematic comparison study indicated that the results from our exact-controllability theory are consistent with those from the structural-controllability theory for cases where both frameworks are applicable. Application of our

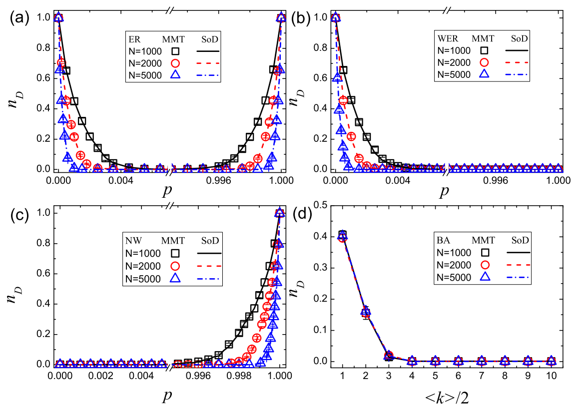

Figure 5: **Exact controllability of unweighted networks**. Exact controllability measure n_D, the number of driver nodes required to control a network, as a function of connecting probability p for (a) unweighted random networks and (b) random networks with random weights assigned to links. (c) n_D versus the probability p of randomly adding links for small-world networks. (d) n_D versus half of average degree $\langle k \rangle/2$ for scale-free networks. All the networks are undirected and their coupling matrices are symmetric. The data points are obtained from the maximum multiplicity theory (MMT) and the error bars denote the standard deviations, each from 20 independent realizations. The curves (SoD) are the theoretical predictions for sparse and dense networks, respectively. The representative network sizes used are $N = 1000$, 2000 and 5000.

framework also revealed a number of phenomena that cannot be uncovered by the structural-controllability framework. For example, we found that for random and small-world networks with identical link weights, the measure of controllability is a non-monotonic function of link density with largest controllability occurring in the intermediate region. For highly sparse or dense networks, the former being ubiquitous in real-world systems, the exact-controllability theory can be greatly simplified, leading to an efficient computational paradigm in terms solely of the rank of the network matrix. Some representative results are shown in Fig. 5.

Our exact-controllability framework can have broader scope of applications than the structural controllability framework. For example, if the weights of partial links are available, our framework will offer better measurement of controllability by setting the weights of other unavailable links to be random parameters, namely, partial structural matrix. Our framework is also valid for undirected networks, where the structural matrix assumption is slightly violated because of the network symmetry. Our framework is as well applicable to networks full of self-loops with identical or distinct weights. Furthermore, investigating exact controllability is important for achieving actual control and predicting control energy, especially in man-made networks. Our exact-controllability theory as an alternative to the structural-controllability theory then offered deeper understanding of our ability to control complex networked systems.

Details of this work can be found in

- Z.-Z. Yuan, C. Zhao, Z.-R. Di, W.-X. Wang, and Y.-C. Lai, “Exact controllability of complex networks,” *Nature Communications* **4**, 2447, 1-9 (2013).

3.14 Emergence of scaling in human-interest dynamics

A fundamental feature of the human society is that its individuals possess all kinds of interests, the driving force of many human behaviors. Some interests may last for a lifetime while others can fade away in short time. From time to time our interests also change. In the modern society that we live in, all kinds of attractions and temptations emerge and disappear on a daily basis. Does this mean that the evolution of our interest is mostly random? Or are there intrinsic dynamical rules that govern how human interests evolve with time? To answer these questions was deemed to be extremely difficult, due to the lack of appropriate means to characterize human mind and to measure quantitatively how it changes with time. Yet the questions are fundamental in science, and any revelation of the dynamics of human interest may have significant applications in commerce, medical sciences, and even defense. In particular, in commerce, adequate knowledge of customer interests and how they change with time are key to the success of many businesses as such knowledge can be of tremendous value to advertisement design and product promotion. In psychiatry, a good understanding of patients' interests may help generate accurate diagnosis and devise effective therapeutic approaches. In defense, timely and reliable assessment of certain group or individuals' interests and their time evolution can help predict the group or individuals' possible future behaviors and actions. Apparently, all these rely on human-interest dynamics' being not completely random.

There had been efforts in modeling and understanding human behaviors that are essential to many social and economical phenomena, with significant applications in areas ranging from resource allocation and transportation control to epidemic prediction and personal recommendation. The pursuit had been facilitated greatly by the advances in information technology, especially by the availability of massive Internet data and resources. However, to probe into human-interest dynamics is more challenging, due to the difficulty in characterizing human interests and traditional lack of data sets from which the underlying dynamical processes may be deduced. In recent years "big data" sets, such as those from e-commerce or mobile-phone communications, become commonly available, making it possible to quantify human interests and to infer their intrinsic dynamics. As a branch of the science of "Big Data," the field of human-interest dynamics is at its infancy.

A viable approach to probing into human-interest dynamics is to use data analysis as a getaway to uncover various phenomena and possible scaling laws. Guided by this principle, we explored two e-commerce data sets (Douban, Taobao) and one communication data set [Mobile-Phone Reading (MPR)], and focused on three issues: statistical distribution of the time that an interest lasts, distribution of the return time to revisit a particular interest, and interest ranking and transition. Considering the large number of factors that can affect human interest, such as the specific activity contents and distractions of the individual's attention, it seems plausible that the underlying dynamics be completely random. Indeed, a widely used assumption is that of Markovian type of dynamics for individuals' online behaviors, in which an online user's next action depends not on his/her history of interests but on the current interest only. However, there had been evidence of deviations from the Markovian dynamics. Our systematic analysis of the three big data sets revealed an unequivocal signature of the power-law scaling behavior characteristic of non-equilibrium complex systems and, consequently, indicated the existence of intrinsic dynamical rules governing the human-interest dynamics. Based on the empirical analysis, we identified three basic ingredients underlying the dynamics: preferential return, inertia effect and exploration. A mathematical model incorporating these ingredients was then developed to account for the observed power-law scaling behaviors. Our study represented the first systematic attempt to probe into the dynamics of human interest, and we expect our finding and model to have broad applications.

As a representative result, we describe here the power-scaling behavior of interest interval l that we uncovered from our "big" data sets. We used categories to characterize an individual's interests, which can be, for example, music, books and movies on *Douban*, clothing, footwear, and toys in *Taobao*, love stories and science fictions on *MPR*, and so on. Figure 6(a) shows, for a typical individual on *Douban*, the distribution $P(l)$ of l visiting different interest categories, which exhibits a power-law scaling: $P(l) \sim l^{-\alpha}$. The long tail associated with the power-law scaling indicates that the individual tends to spend an abnormally long time visiting certain interests during browsing. Similar scaling behaviors have been found for users on *Taobao* and *MPR*, as shown in Figs. 6(b) and 6(c), respectively. A typical sequence that the values of l corresponding to an identical interest appear is shown in Fig. 6(d), where we observe a highly non-uniform behavior in the values of l, giving rise to the power-law distribution in Fig. 6(a). We examined many individuals from the

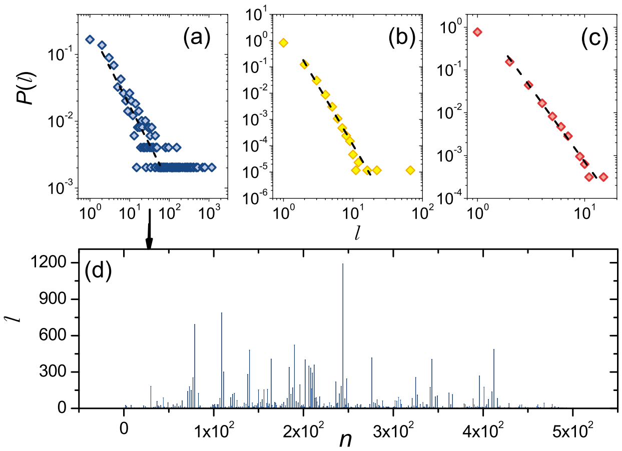

Figure 6: **Distribution of interest-dwelling time**. (a-c) Probability distributions $P(l)$ of the time interval l of consecutive visits to the same interest for three representative individuals, each from one of the three data sets (*Douban*, *Taobao*, and *MPR*), where the numbers of interests are 3, 24, and 44, respectively. The numbers of clicks (N_a) for the three cases are are 18396, 106571, and 4398, respectively. The three distributions can be fitted as power-law $P(l) \sim l^{-\alpha}$, with exponents $\alpha \approx 1.16, 4.02$ and 3.35, respectively. Panel (d) shows the various values of l as they appear with time, where n is the event index (an integer variable).

three data sets, and found similar power-law behaviors. In fact, the distribution of l for *all* users from any particular data set exhibits a robust power-law scaling. The power-law scaling observed for all cases implies substantial derivation of the human-interest dynamics from that of the Markovian process (associated with the transition probability matrix for interests) for which the scaling of l would be exponential.

The scaling laws uncovered from data and the dynamical model developed accordingly can be applied to addressing significant problems ranging from human-behavior prediction and the design of search algorithms to controlling spreading dynamics. As a demonstration, we quantified the degree of predictability of user-behavior patterns underlying the three data sets by using the statistical measures of entropy and Fano inequality, with the result that such patterns are in fact quite predictable, despite the apparent randomness in the human-interest dynamics.

Details of this work can be found in

- Z.-D. Zhao, Z.-M. Yang, Z.-K. Zhang, T. Zhou, Z.-G. Huang, and Y.-C. Lai, "Emergence of scaling in human-interest dynamics," *Nature Scientific Reports* **3**, 3472, 1-7 (2013).

3.15 Robustness of chimera states in complex dynamical systems

The collective dynamics of complex systems are often multifold and much more complicated than the dynamics of individual oscillators. For example, when a large number of oscillators, each possessing very simple dynamics, are coupled together, the collective behaviors of all the oscillators can be highly nontrivial. In the classic Kuramoto network, each oscillator is coupled with every other oscillator - the configuration of a globally coupled network. Each individual oscillator is a simple rotation of certain frequency, and the dynamics of the oscillators differ only in their frequencies. The coupling function is also a simple mathematical function, such as a sinusoidal type of function. For relatively weak coupling the motions of the oscillators are incoherent, due to the heterogeneity in their frequencies, but as the coupling parameter increases through a critical value, coherence can emerge and persist in the form of partial or complete synchronization. There exists a large body of literature on synchronization in the Kuramoto network, due to its relevance to many physical, chemical, and biological phenomena.

While the emergence of synchronous behavior as the coupling is strengthened is intuitively reasonable and anticipated in any coupled oscillator network, complex systems often present us with unexpected and sometimes quite surprising phenomena. A striking example is the occurrence of chimera state in non-locally coupled networks of *identical* oscillators, where different subsets of the oscillators can exhibit completely distinct dynamical behaviors. For example, for a simple form of chimera state, there are two distinct types of behavior among all oscillators in the network: one group of oscillators is nearly perfectly synchronous but the oscillators in the complementary group are completely incoherent. These two types of behaviors emerge as one state of the networked system, in contrast to the phenomenon of multiple coexisting attractors in nonlinear dynamical systems, each with its own basin of attraction. In such a system, while the attractors coexist in the phase space, starting from a single initial condition the system approaches asymptotically to only one attractor of certain characteristics, which can be a stable fixed point, a limit cycle, a quasiperiodic state, or even a chaotic attractor, but from the same initial condition the system cannot simultaneously possess more than one of these traits. Signatures of chimera states were first observed from the spatiotemporal evolution of a system of coupled nonlinear oscillators and the phenomenon was named "domain-like spatial structure." Chimera states in highly regular and non-locally coupled networks of identical oscillators are thus a quite remarkable type of collective dynamics. We note that nonlocal coupling is relevant to physical systems such as the Josephson-junction arrays and to chemical oscillators as well.

The paradigmatic setting in which chimera states had been studied theoretically and computationally is that of non-locally coupled phase oscillators. A fundamental question was how robust chimera states are with respect to perturbations. That is, when the system details deviate from those of the paradigmatic setting or when noise is present, can chimera states still emerge and sustain? In this regard, the issue of noise was successfully addressed, as chimera states had been experimentally observed in a chemical and an optical systems that are intrinsically noisy. An outstanding issue is then how random perturbations to the network structure affect the chimera states. We addressed this *structural robustness* issue that is fundamental to our understanding of chimera states. In particular, starting from the standard setting of a non-locally coupled array of identical phase oscillators, we removed links systematically but randomly according to the removal probability p and investigated whether and to what extent chimera states can persist as p is increased from zero. For a fixed value of p, for an infinite network there are an infinite number of possible configurations. For a realistic network of finite size, the number of configurations can still be extremely large. Due to randomness in the network structure, the persistence of chimera states can be characterized but in a statistical sense. In particular, given p, certain fraction of the network configurations would permit chimera states, while others would not. One can thus define a probability for chimera states, denoted by $F(p)$, where $F(p) \to 1$ for $p \to 0$ and in general we expect $F(p)$ to be a decreasing function of p. Our extensive computations revealed that chimera states can persist for a range of p values in the sense that $F(p)$ maintains values close to unity even when p is appreciably away from zero, strongly suggesting that the exotic dynamical states are robust with respect to random structural perturbations to the underlying network. We then resorted to two independent theoretical approaches, one based on self-consistency and another based on the spectral theory for collective dynamics on networks. Both gave results that are consistent with those from direct numerical computations. A surprising finding is that, even for relatively large values of p for which a large number of links have been removed and chimera state is deemed unlikely, the division of oscillators into coherent and incoherent groups persists. The commonly recognized chimera state, which occurs for smaller values of p, is nothing but a particular case in which the coherent oscillators happen to be synchronized or phase-locked. Some representative results are shown in Fig. 7.

The phenomenon that we uncovered is rather striking: regardless of whether chimera state can emerge, the system exhibits a general breathing pattern in its spatiotemporal evolution. Associated with such a pattern, the oscillators in the system can be qualitatively classified into two groups: one group of high coherence and another of weak coherence. The particular breathing pattern stipulates that this division holds even for large link-removal probability where chimera state is ruled out. The implication is that the breathing pattern in the spatiotemporal evolution of the system is general and robust, and chimera state is a particular phenomenon where the oscillators in the highly coherent group happen to be phase synchronized. Our work thus provided deeper insights into the dynamical origin of chimera state, a phenomenon of continuous

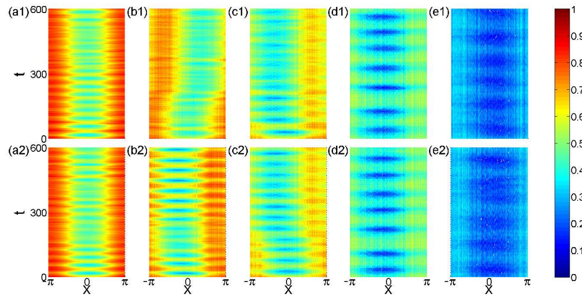

Figure 7: **Spatiotemporal evolution of the order parameter associated with chimera states.** Contour plots representing spatiotemporal evolution of the order parameter for five values of the link removal probability p (from left to right: 0, 0.1, 0.2, 0.4, and 0.6). The five patterns in the top row are obtained by the PDE in the continuum limit, and the corresponding patterns in the bottom row are from direct numerical calculations of the original dynamical system.

interest and subject to intense recent investigation.

More details of this work can be found in

- N. Yao, Z.-G. Huang, Y.-C. Lai, and Z.-G. Zheng, "Robustness of chimera states in complex dynamical systems," *Nature Scientific Reports* **3**, 3522, 1-8 (2013).

3.16 Uncovering hidden nodes in complex networks in the presence of noise

The difficulty to develop effective methods for detecting hidden nodes in complex networks is compounded by the fact that the indirect information on which any method of detecting hidden objects relies can be subtle and sensitive to changes in the system or in the environment. In particular, in realistic situations noise and random disturbances are present. It is conceivable that the "indirect" information can be mixed up with that due to noise or be severely contaminated. The presence of noise thus poses a serious challenge to detecting hidden nodes, and some effective "noise-mitigation" method must be developed.

To formulate the problem in a concrete way and to gain insights into the development of a general methodology, we noted that the basic principle underlying the detection of hidden objects is that their existence typically leads to "anomalies" in the quantities that can be calculated or deduced from observation. Simultaneously, noise, especially local random disturbances applied at the nodal level, can also lead to large variance in these quantities. This is so because, a hidden node is typically connected to a few nodes in the network that are accessible to the external world, and a noise source acting on a particular node in the network may also be regarded as some kind of hidden object. Thus, the key to any detection methodology is to identify and *distinguish* the effects of hidden nodes on measures for detection from those due to *local* noise sources.

In our work, we focused on complex networks and developed a general method to differentiate hidden nodes from local noise sources. This problem is intimately related to the works on reverse engineering of complex networks, where the goal is to uncover the full topology of the network based on measured time series. Our method was based on compressive sensing to detect hidden nodes in the absence of noise sources. To explain our method in a concrete setting, here we use the network configuration shown schematically in Fig. 8, where there are 20 nodes, the couplings among the nodes are weighted, and the entire network is in a noisy environment, but a number of nodes also receive relatively strong random driving. We assume an oscillator network so that the nodal dynamics are described by nonlinear differential equations, and that time series can be measured simultaneously from all nodes in the network except one, labeled as #20, which is a hidden node. The tasks of ascertaining the presence and locating the position of the hidden node are equivalent to identifying its immediate neighbors, which are nodes #3 and #7 in Fig. 8. Note that, in order to be able to detect the hidden node based on information from its neighboring nodes, the interactions between the hidden node and its neighbors must be directional from the former to the latter or be bidirectional.

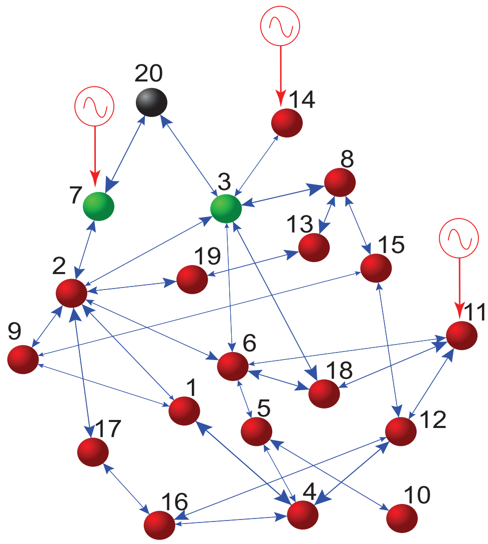

Figure 8: **Detection of hidden node in the presence of noise.** An example of a complex network with a hidden node. Time series from all nodes except hidden node #20 can be measured, which can be detected when its immediate neighbors, nodes #3 and #7 are unambiguously identified. Nodes #7, #11, and #14 are driven by local noise sources.

Otherwise, if the coupling is solely from the neighbors to the hidden node, the dynamics of the neighboring nodes will not be affected by the hidden node and, consequently, time series from the neighboring nodes will contain absolutely no information about the hidden node, which is therefore undetectable. The action of local noise source on a node is naturally directional, i.e., from the source to the node.

We had demonstrated that, when the compressive-sensing paradigm is applied to uncovering the network topology, the predicted linkages associated with nodes #3 and #7 are typically anomalously dense, and this piece of information is basically what is needed to identify them as the neighboring nodes of the hidden node. In addition, when different segments of measurement data are used to reconstruct the coupling weights for these two nodes, the reconstructed weights associated with these two nodes exhibit significantly larger variances than those associated with other nodes. However, the predicted linkages associated with the nodes driven by local noise sources can exhibit behaviors similar to those due to the hidden nodes, leading to uncertainty in the detection of the hidden node. To address this critical issue is essential to developing algorithms for real-world applications. Our main idea was to exploit the principle of *differential signal* to study the behavior of the predicted link weights as a function of the data used in the reconstruction. Due to the advantage of compressive sensing, the required data amount can be quite small and, hence, even if our method requires systematic increase in the data amount, it will still be reasonably small. We argued and demonstrated that, when the various ratios of the predicted weights associated with all pairs of links between the possible neighboring nodes and the hidden node are examined, those associated with the hidden nodes and nodes under strong local noise should show characteristically distinct behaviors, rendering unambiguous identification of the neighboring nodes of the hidden node. Any such ratio is essentially a kind of differential signal, because it is defined with respect to a pair of edges. Representative results are shown in Fig. 9.

Detecting hidden nodes in complex networks with *a priori* unknown nodal dynamics, topology, and coupling weights has vast application potential, as in social and biological networks. Inferring the existence of hidden node in the presence of local random perturbations is an extremely challenging problem. Our efforts represented a step forward in this area of research, where much further work is needed.

More details of this work can be found in

- R.-Q. Su, Y.-C. Lai, X. Wang, and Y.-H. Do, "Uncovering hidden nodes in complex network in the presence of noise," *Nature Scientific Reports* **4**, 3944, 1-7 (2014).

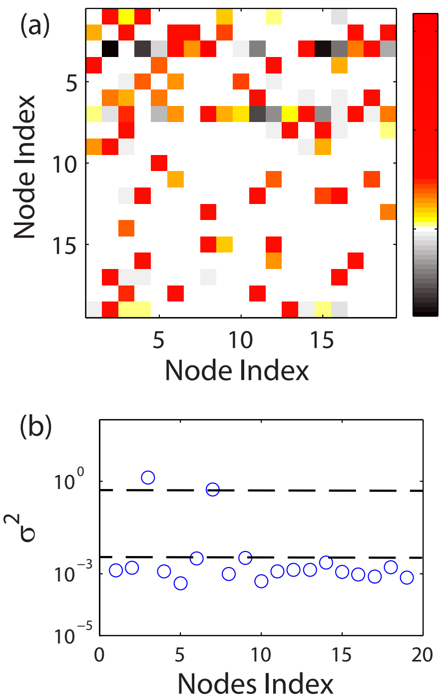

Figure 9: **Detection of hidden nodes in the presence of noise.** For the network in Fig. (8), (a) predicted coupling matrix for all nodes except node #20. Time series from nodes #1 to #19 are available, while node #20 is hidden. The predicted weights are indicated by color coding and the amount of data used is $R_m = 0.7$. The abnormally dense patterns in the 3rd and 7th rows suggest that nodes #3 and #7 are the immediate neighbors of the hidden node. (b) Variance σ^2 of the predicted coefficients for all accessible nodes, which is calculated using 20 independent reconstructions based on different segments of the data. The variances associated with nodes #3 and #7 are apparently much larger than those with the other nodes, confirming that these are the neighboring nodes of the hidden node. There is a definite gap between the values of the variance associated with neighboring and non-neighboring nodes of the hidden node, as indicated by the two horizontal dashed lines in (b). When the local noise sources are applied to node #7, #11 and #14, these there nodes have similar dense bars in (a) and large variances in (b).

3.17 Reconstructing propagation networks with natural diversity and identifying hidden source

An important class of collective dynamics is epidemic spreading and information diffusion in the human society or on computer networks. The past decades have witnessed severe epidemic outbreaks at the global scale due to the mutation of virus, including SARS, H5N1, and H7N9 in eastern China. Our goal was to reconstruct the networks hosting the spreading process and identify the source of spreading using limited measurements. This is especially challenging due to (1) difficulty in predicting and monitoring mutations of deadly virus and (2) absence of epidemic threshold in heterogeneous networks. Another example is rumor propagation in the online virtual communities, which can cause financial loss or even social instabilities, such as the 2011 irrational and panicked acquisition of salt in southeast Asian countries caused by the nuclear leak in Japan. In this regard, identifying the propagation network for controlling the dynamics is of great interest. Another significant challenge in reconstructing a spreading network lies in the nature of the available time series: they are polarized, despite stochastic spreading among nodes. Indeed, the link pattern and the probability of infection are encrypted in the binary status of individuals, infected or not, analogous to the collapse of wave function to one associated with some discrete quantum state induced by observation in quantum mechanics.

There had been previous efforts in addressing the inverse problem of some special types of complex propagation networks. In particular, for diffusion process originated from a single source, the routes of diffusion from the source constitute a tree-like structure. If information about the early stage of the spreading dynamics is available, it would be feasible to decode all branches that reveal the connections from the source to its neighbors, and then to their neighbors, and so on. Taking into account the time delays in the diffusion process enables a straightforward inference of the source in a complex network through enumerating all possible hierarchical trees. However, if no immediate information about the diffusion is available, the tree-structure based inference method is inapplicable, and the problem of network reconstruction and locating the source becomes extremely challenging, hindering control of diffusion and delivery of immunization.

The loss of knowledge about the source is common in real situations. For example, passengers on an international flight can carry a highly contagious disease, making certain airports the immediate neighbors of the hidden source, which would be difficult to trace. In another example, the source could be migratory birds coming from other countries or continents. A general data-driven approach, applicable in such scenarios, was lacking.

Sponsored by AFOSR, we developed a general theoretical framework to reconstruct complex propagation networks from time series based on compressed sensing theory (CST). Due to the striking characteristics of CST such as the extremely low data requirement and rigorous guarantee of convergence to optimal solutions, our framework is highly efficient and accurate. However, casting the inverse problem into the CST framework is highly nontrivial. Although CST has been used to uncover the nodal interaction patterns for coupled oscillator networks or evolutionary games from time series, the dynamics of epidemic propagation is typically highly stochastic with, for example, binary time series, rendering inapplicable previous CST-based formulations. Further, despite the use of alternative sparsity enforcing regularizers and convex optimization to infer networks, CST had not been applied to reconstructing propagation networks, especially when the available time series are binary. Our main accomplishment was then the development of a scheme to implement the highly nontrivial transformation associated with the spreading dynamics in the paradigm of CST. Without loss of generality, we employed two prototypical models of epidemic spreading: classic susceptible-infected-susceptible (SIS) dynamics and contact processes (CP), on both model and real-world (empirical) networks. Inhomogeneous infection and recovery rates as representative characteristics of the natural diversity are incorporated into the diffusion dynamics to better mimic the real-world situation. We assumed that only binary time series can be measured, which characterize the status of any node, infected or susceptible, at any time after the outbreak of the epidemic. The source that triggers the spreading process is assumed to be externally inaccessible (hidden). In fact, one may not even realize its existence from available time series. Our method enabled, based on relatively small amounts of data, a full reconstruction of the epidemic spreading network with nodal diversity and successful identification of the immediate neighboring nodes of the hidden source (thereby ascertaining its existence and uniquely specifying its connections to nodes in the network). The framework was validated with respect to different amounts of data generated from various combinations of the network structures and dynamical processes. High accuracy, high efficiency and applicability in a strongly stochastic environment with measurement noise and missing information are the most striking characteristics of our framework. Thus broad applications can be expected in addressing significant problems such as targeted control of disease and rumor spreading.

More details of this work can be found in

- Z.-S. Shen, W.-X. Wang, Y. Fan, Z.-R. Di, and Y.-C. Lai, "Reconstructing propagation networks with natural diversity and identifying hidden source," *Nature Communications* **5**, 4323, 1-10 (2014).

3.18 Spatiotemporal patterns and predictability of cyberattacks

Highly networked communication and information infrastructures built via various state-of-the-art technologies play crucial roles in modern economic, social, military, and political activities. However, such sophisticated infrastructures are facing more and more severe security challenges on the global scale. Earlier theoretical works focused on understanding the complex topologies of the Internet and on the likelihood of large scale failures caused by node removal in complex networks. Recent years have witnessed tremendous efforts devoted to mitigating and coping with increasing cybersecurity threats. For example, attack graphs were invented to analyze the overall network vulnerability and to generate a global view of network security against attacks. By deploying network sensors at particular points in the Internet, monitoring systems were built to detect cyberthreats and statistically analyze the time, sources, and the types of attacks, and various visualization methods were developed to better understand the result of the detection and analysis. Quite recently, a genetic epidemiology approach to cybersecurity was proposed to understand the factors that determine the likelihood that individual computers are compromised, and the general concept of cybersecurity dynamics was introduced.

Attack traffic analysis were mainly done in the field of Intrusion Detection System (IDS), the cyberspace's equivalent to the burglar alarm. IDS has become one of the fundamental technologies for network security. There are three approaches to building an IDS: (1) signature or misuse detection, (2) anomaly detection, and (3) hybrid or compound detection. Specifically, signature detection technique is based on a predefined set of known attack signatures obtained from security experts. The system observes the activities of subjects and alarms if their behaviors match the malicious ones in the attack signature set. Both host-based and network-based detection systems were developed. Anomaly detection technique is based on machine learning methodologies, such as system-call based sequence analysis, Bayesian networks, principal component analysis, and Markov models. The IDS monitoring capability can be improved by taking a hybrid approach that combines both signature and anomaly detection strategies. All these methods are often based on data packet payload inspection and thus are difficult to perform for high speed networks. Another limitation of these approaches is the assumption that either the attacks are well defined (i.e., signatures) or the normal behaviors are well defined (so are the abnormal behaviors). There had been a growing interest in flow-based intrusion detection technologies, by which communication patterns within the network are analyzed, instead of the contents of individual packets. Interestingly, a quite recent study analyzing the data obtained from the host IDSs revealed strong associations between the network services running on the host and the specific types of threats to which it is susceptible. Making use of the plan recognition method in artificial intelligence, one can predict the attack plan from the IDS alert information. Utilizing virtual or physical networks to test these IDS techniques can be costly and time consuming, hence, as an alternative, simulation modeling approaches were developed to represent computer networks and IDS to efficiently simulate cyberattack scenarios. As botnets have become a major threat in cyberspace, cyberattack traffic patterns had also been used to understand botnet's Command-and-Control strategies.

We uncovered the existence of intrinsic spatiotemporal patterns underlying cyberattacks and addressed the important question of whether certain such attacks may be predicted or anticipated in advance. The overwhelming complexity of the modern cyberspace would suggest complete randomness in the distribution of cyberattacks and, as a result, the intuitive expectation is that attackers' behaviors are random and attacks are unpredictable. However, our discovery of the spatiotemporal patterns and quantitative characterization of the predictability of these patterns suggested the otherwise. In particular, distinct from previous works on cyberattack analysis, our efforts concentrate on analyzing the *macroscopic* properties of the attack traffic flows using a data set of cyberattacks available to us. Especially, the data set recorded attacks on 491 consecutive victim IP addresses (sensors) in 18 days. The IP addresses can thus be regarded, approximately, as a variable in space. An attack is regarded as an event occurring in both space and time, and we can speak of events in spatiotemporal dimensions. This is much more comprehensive than the analysis of the individual time series obtained from sampled IP addresses or the time series obtained by treating the IP addresses as a whole. Our results revealed, for the first time, that robust macroscopic patterns exist in the seemingly random cyberspace: majority of the attacks are governed by a few very limited number of patterns, indicating that cyberattacks are mainly committed by a few types of major attackers, each with unique spatiotemporal characteristics. More specifically, the patterns can be divided into two types: deterministic and stochastic. The emergence of deterministic patterns implies predictability, which can potentially be exploited to anticipate certain types of attacks to achieve greater cybersecurity. We characterized the predictability of attack frequency time series based on information entropy. Our results suggested a surprisingly high degree of predictability, especially for the IPs under deterministic attack. Effective algorithms can then be developed to predict the future attack frequencies. We also developed methods to evaluate the inference probability between the attack frequency time series based on series similarity, which may allow us to plant much fewer attack probes into the Internet while still achieving effective monitoring. The stochastic patterns can be quantified using the flux-fluctuation law in statistical and nonlinear physics. Our findings outlined a global picture of how cyberattacks are initiated and distributed into the Internet. This would be of potential value to the development of defense strategies against cyberattacks on a global scale.

More details of this work can be found in

- Y.-Z. Chen, Z.-G. Huang, S.-H. Xu, and Y.-C. Lai, "Spatiotemporal patterns and predictability of cyberattacks," *PLoS ONE*, accepted (to appear in June 2015).

4 Personnel Supported and Theses Supervised by PI

4.1 Personnel Supported

The following people received salaries from the AFOSR Project during various time periods.

- **Faculty**: Ying-Cheng Lai (PI), ISS Chair Professor of Electrical Engineering, Professor of Physics.
- **Post-Doctoral Fellows**: Dr. Wenxu Wang and Dr. Zigang Huang.
- **Ph.D. Students**: Rui Yang (2012), Xuan Ni (2013), Riqi Su (to graduate in 2015), Yuzhong Chen (to graduate in 2016), Lezhi Wang (ongoing).

4.2 Ph.D. graduates who participated in research in the project area

1. Rui Yang, Electrical Engineering, ASU, May 2012. Ph.D. Dissertation: *System reconstruction via compressive sensing, complex-network dynamics, and electronic transport in graphene systems*.
2. Xuan Ni, Electrical Engineering, ASU, May 2013. Ph.D. Dissertation: *Effect of chaos on relativistic quantum tunneling*; **Recipient of 2012-2013 Palais Outstanding Doctoral Student Award, ASU ECEE**.

5 Interactions/Transitions

5.1 Collaboration with AFOSR scientist

Dr. Vassilios Kovanis from the Air Force Research Laboratory at Wright Patterson Air Force Base, on compressive-sensing based identification of complex dynamical systems and networks.

5.2 Invited talks on topics derived from the project

During the project period, PI gave the following invited plenary talks, seminars, and colloquia on various topics derived from AFOSR sponsored research.

1. "Predicting complex networks and dynamical systems based on time series," 2010 NIMS (National Institute for Mathematical Sciences) International Workshop on Applied Dynamical Systems, Daejeon, South Korea; December 8, 2010.

2. “Predicting complex networks and dynamical systems based on time series,” Plenary talk, The 1st International Symposium on Innovative Mathematical Modelling, University of Tokyo, Japan; March 2, 2011.
3. “Catastrophic dynamics on complex networks: prediction and control,” Invited talk, NSF Workshop on Building Engineering Complex Systems, Arlington, VA; March 29, 2011.
4. “Uncovering complex-network topologies and dynamical systems based on time series,” Invited talk, XXXI European Dynamics Days Conference, University of Oldenburg, Germany; September 13, 2011.
5. “Time-series based prediction of nonlinear dynamical systems and complex networks,” Seminar, Center for Biological Physics, Arizona State University; September 21, 2011.
6. “Reverse engineering of nonlinear dynamical systems and complex networks - a compressive-sensing based approach,” Plenary talk, International Conference on Modeling Life Sciences, Fudan University, Shanghai, China; September 26, 2011.
7. “Reverse engineering of nonlinear dynamical systems and complex networks,” Colloquium, Department of Physics, Eastern China Normal University, Shanghai, China; September 27, 2011.
8. “Transient chaos,” Plenary talk, The Fourth International Workshop on Chaos-Fractals: Theory and Applications, Hangzhou Dianzi University, Hangzhou, China; October 22, 2011.
9. “Introduction to transient chaos,” Undergraduate colloquium, Department of Physics, Lanzhou University, Lanzhou, China; October 25, 2011.
10. “Transient Chaos,” Plenary talk, International Workshop on Anomalous Statistics, Generalized Entropies, and Information Geometry, Nara Women’s University, Nara, Japan; March 7, 2012.
11. “Uncovering complex-network topologies and dynamical systems based on compressive sensing,” Plenary lecture, International Symposium on Compressed Sensing: Theory and Applications, Tianjin University, Tianjin, China; June 9, 2012.
12. “Transient Chaos,” Colloquium, School of Electrical Engineering and Automation, Tianjin University, Tianjin, China; June 10, 2012.
13. “Controlling complex networks,” Plenary talk, 5th Shanghai International Symposium on Nonlinear Science and Applications, Fudan University, Shanghai, China; June 28, 2012.
14. “Complex networks: controllability and control of collective dynamics,” Graduate Colloquium, School of Physics, Lanzhou University, Lanzhou, China; July 4, 2012.
15. “Complex networks: controllability and control of collective dynamics,” Colloquium, School of Electrical Engineering and Automation, Xi’an University of Technology, Xi’an, China; July 9, 2012.
16. “Transient chaos,” Plenary talk, Dynamics Days Asia Pacific 7 - The 7th International Conference on Nonlinear Science & the 11th Taiwan International Symposium on Statistical Physics, Academia Sinica, Taipei, Taiwan; August 6, 2012.
17. “Research on nonlinear dynamics and complex systems for applied mathematics - a vision,” Distinguished University Lecture (hosted by the President of the University), Kyungpook National University, Daegu, South Korea; September 11, 2012.
18. “Predicting dynamical systems and complex networks via compressive sensing,” Colloquium, Department of Applied Mathematics, Ulsan National Institute of Science and Technology, Ulsan, South Korea; November 14, 2012.

19. “Predicting complex dynamical systems via compressive sensing,” Seminar, Department of Mathematics, Kyungpook National University, Daegu, South Korea; December 6, 2012.
20. “Reverse engineering of nonlinear dynamical systems and complex networks,” Invited talk, NSF Workshop on Building Engineered Complex Systems, Arlington, VA; January 24, 2013.
21. “Recent advances in complex networks,” Graduate Seminar Series (fifteen 90-minute seminars), on Sabbatical Leave at Kyungpook National University, South Korea; March 4 - June 10, 2013.
22. “Hidden nodes, extreme events, human-interest dynamics, and quantum entanglement,” Invited talk, Invited-Only ARO Workshop on Information in Complex Dynamical Systems, Burlington, VM; July 18, 2013.
23. “Nonlinear dynamics and complex systems - a mathematical paradigm for cutting-edge, interdisciplinary research,” Invited Keynote talk, IBS (Institute for Basic Sciences) International Symposium on “Towards a mathematical theory of nonlinear dynamical and complex systems,” Seoul, South Korea; August 1, 2013.
24. “Controlling nonlinear dynamics on complex networks,” Invited talk, Satellite Symposium of NetSci 2014 (International School and Conference on Network Science), Berkeley, California; June 2, 2014.
25. “Controlling nonlinear dynamics on complex networks,” Plenary talk, The 6th Shanghai International Symposium on Nonlinear Sciences and Applications, Fudan University, Shanghai, China; June 29, 2014.
26. “Data based reconstruction of complex networks and energy optimization,” Keynote lecture, North America-East Asia Workshop on Big Data Analytics for Infrastructure and Building Sustainability and Resilience (IBSR) Research, Beijing, China; September 19, 2014.
27. “Transient dynamics in nonlinear and complex systems,” Invited talk, ARO Workshop on Cyber Security Dynamics, University of North Carolina, Chapel Hill; September 23, 2014.
28. “Data based reconstruction of complex dynamical systems,” Colloquium, University of Missouri, Columbia, MO; October 30, 2014.
29. “Nonlinear dynamics and complex systems - a paradigm for cutting-edge, interdisciplinary research,” Physics Colloquium, Shanxi Normal University, Xi’an, China; March 11, 2015.

6 Past Honors

1. NSF Faculty Career Award, 1997.
2. Air Force PECASE, 1997.
3. Election as a Fellow of the American Physical Society, 1999. Citation: *For his many contributions to the fundamentals of nonlinear dynamics and chaos*.
4. NSF ITR Award, 2003.
5. Outstanding Referee Award, American Physical Society, 2008.

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

Ying-Cheng.Lai@asu.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

1-480-965-6668

Organization / Institution name

Arizona State University

Grant/Contract Title
The full title of the funded effort.

PREDICTING AND CONTROLLING COMPLEX NETWORKS

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-10-1-0083

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

Ying-Cheng Lai

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. Tristan Nguyen

Reporting Period Start Date

04/01/2010

Reporting Period End Date

03/31/2015

Abstract

The principal Objective of the project was to develop methods to predict and control complex networks. For prediction, a number of methods were articulated and tested to uncover the structures and topologies of complex networks as well as various dynamical processes on the networks based solely on time series data or measured signals. A compressive sensing based framework for network and nonlinear dynamical systems reconstruction was pioneered. For control, key issues including linear controllability of complex networks, control energy, control of collective dynamics, and control of nonlinear dynamics on complex networks were addressed. A number of new phenomena in complex dynamical systems were uncovered and understood, and computational paradigms were established for prediction and control. The AFOSR project resulted in 50 refereed-journal papers, including papers in high-impact journals such as Physical Review Letters, Nature Communications, and Physical Review X. The AFOSR support provided PI with the opportunity to supervise a number of PhD students: two graduated, one to graduate in 2015, and two ongoing. PI gave about 30 plenary lectures, seminars, and colloquiums all over the world on predicting and controlling complex networks.

Specific accomplishments include (1) uncovering the full topology of oscillator networks, (2) cascading failures and the emergence of cooperation in evolutionary-game based models of social and economic

networks, (3) information explosion on complex networks and control, (4) pattern formation, synchronization and outbreak of biodiversity in cyclically competing games, (5) predicting catastrophes in nonlinear dynamical systems by compressive sensing, (6) time-series based prediction of complex oscillator networks via compressive sensing, (7) reconstruction of social networks based on evolutionary-game data via compressive sensing, (8) optimizing controllability of complex networks by minimum structural perturbations, (9) detecting hidden nodes in complex networks from time series based on compressive sensing, (10) Forecasting synchronizability of complex networks from data, (11) emergence of grouping in multi-resource minority game dynamics, (12) optimizing cooperation on complex networks in the presence of failure, (13) exact controllability of complex networks, (14) emergence of scaling in human interest dynamics, (15) robustness of chimera states in complex dynamical systems, (16) uncovering hidden nodes in complex networks in the presence of noise, (17) reconstructing propagation networks with natural diversity and identifying hidden source, and (18) spatiotemporal patterns and predictability of cyberattacks.

Distribution Statement

This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

SF298_FA9550-10-1-0083.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

Final_Report_FA9550-10-1-0083_Lai.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

1. J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, “Noise bridges dynamical correlation and topology in coupled oscillator networks,” Physical Review Letters 104, 058701, 1-4 (2010).

2. H.-J. Shi, R. Yang, W.-X. Wang, and Y.-C. Lai, “Basins of attraction for species extinction and coexistence in spatial rock-paper-scissors games,” Physical Review E (Rapid Communications) 81, 030901(R), 1-4 (2010).

3. W.-X. Wang, Y.-C. Lai, and C. Grebogi, “Effect of epidemic spreading on species coexistence in spatial games,” Physical Review E 81, 046113, 1-4 (2010). One figure from this paper was selected for “Kaleidoscope” of PRE.

4. R. Yang, W.-X. Wang, Y.-C. Lai, and C. Grebogi, “Role of intraspecific competition in the coexistence of mobile populations in spatially extended ecosystems,” Chaos 20, 023113, 1-6 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the June 1, 2010 issue (http://www.vjbio.org).

5. X. Ni, W.-X. Wang, Y.-C. Lai, and C. Grebogi, “Cyclic competition of mobile species on continuous space: pattern formation and coexistence,” Physical Review E 82, 066211, 1-8 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the January 1, 2011 issue (http://www.vjbio.org).

6. X. Ni, R. Yang, W.-X. Wang, Y.-C. Lai, and C. Grebogi, “Basins of coexistence and extinction in spatially extended ecosystems of cyclically competing species,” Chaos 20, 045116, 1-8 (2010). This work was selected by the Virtual Journal of Biological Physics Research for the January 1, 2011 issue

(http://www.vjbio.org).

7. W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, “Predicting catastrophes in nonlinear dynamical systems by compressive sensing,” Physical Review Letters 106, 154101, 1-4 (2011).

8. W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, “Pattern formation, synchronization and outbreak of biodiversity in cyclically competing games,” Physical Review E 83, 011917, 1-9 (2011).

9. R.-R. Liu, W.-X. Wang, Y.-C. Lai, G.-R. Chen, and B.-H. Wang, “Optimal convergence in naming game with geography-based negotiation on small-world networks,” Physics Letters A 375, 363-367 (2011).

10. H.-X. Yang, W.-X. Wang, Y.-B. Xie, Y.-C. Lai, and B.-W. Wang, “Transportation dynamics on net- works of mobile agents,” Physical Review E 83, 016102, 1-5 (2011).

11. L.-L. Jiang, M. Perc, W.-X. Wang, Y.-C. Lai and B.-H. Wang, “Impact of link deletions on public cooperation in scale-free networks,” Europhysics Letters 93, 40001, 1-6 (2011).

12. W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and M. A. F. Harrison, “Time-series based prediction of complex oscillator networks via compressive sensing,” Europhysics Letters 94, 48006, 1-6 (2011).

13. L. Huang and Y.-C. Lai, “Cascading dynamics in complex quantum networks,” Chaos 21, 025107, 1-6 (2011). This work was selected by July 2011 issue of Virtual Journal of Quantum Information (http://www.vjquantuminfo.org).

14. W.-X. Wang, Y.-C. Lai, and D. Armbruster, “Cascading failures and the emergence of cooperation in evolutionary game based models of social and economical networks,” Chaos 21, 033112, 1-12 (2011).

15. H.-X. Yang, W.-X. Wang, Y.-C. Lai, Y.-B. Xie, and B.-H. Wang, “Control of epidemic spreading on complex networks by local traffic dynamics,” Physical Review E (Rapid Communication) 84, 045101(R), 1-4 (2011).

16. W.-X. Wang, Y.-C. Lai, C. Grebogi, and J.-P. Ye, “Network reconstruction based on evolutionary- game data via compressive sensing,” Physical Review X 1, 021021, 1-7 (2011).

17. R.-R. Liu, W.-X. Wang, Y.-C. Lai, and B.-H. Wang, “Cascading dynamics on random networks: crossover in phase transition,” Physical Review E 85, 026110, 1-5 (2012).

18. W.-X. Wang, X. Ni, Y.-C. Lai, and C. Grebogi, “Optimizing controllability of complex networks by small structural perturbations,” Physical Review E 85, 026115, 1-5 (2012).

19. G.-M. Zhu, H.-J. Yang, R. Yang, J. Ren, B. Li, and Y.-C. Lai, “Uncovering evolutionary ages of nodes in complex networks,” European Journal of Physics B 85, 106, 1-6 (2012).

20. G. Yan, J. Ren, Y.-C. Lai, C. H. Lai, and B. Li, “Controlling complex networks - how much energy is needed?” Physical Review Letters 108, 218703, 1-5 (2012).

21. R.-Q. Su, X. Ni, W.-X. Wang, and Y.-C. Lai, “Forecasting synchronizability of complex networks from data,” Physical Review E 85, 056220, 1-11 (2012).

22. H.-X.Yang,W.-X.Wang,Y.-C.Lai,andB.-H.Wang,“Traffic-drivenepidemicspreadingonnetworks of mobile agents,” Europhysics Letters 98, 68003, 1-5 (2012).

23. L.-L. Jiang, W.-X. Wang, Y.-C. Lai, and X. Ni, “Multi-armed spirals and multi-pairs antispirals in spatial

rock-paper-scissors games," Physics Letters A 376, 2292-2297 (2012).

24. R.-Q. Su, W.-X. Wang, and Y.-C. Lai, "Detecting hidden nodes in complex networks from time series," Physical Review E (Rapid Communication) 85, 065201(R), 1-4 (2012).

25. R. Yang, Y.-C. Lai, and C. Grebogi, "Forecasting the future: is it possible for time-varying nonlinear dynamical systems?" Chaos 22, 033119, 1-6 (2012).

26. F. Ricci, R. Tonelli, L. Huang, and Y.-C. Lai, "Onset of chaotic phase synchronization in complex networks of coupled heterogeneous oscillators," Physical Review E 86, 027201, 1-4 (2012).

27. W.-X. Wang, J. Ren, Y.-C. Lai, and B. Li, "Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series," Chaos 22, 033131, 1-8 (2012).

28. Z.-G. Huang, J.-Q. Zhang, J.-W. Dong, L. Huang, and Y.-C. Lai, "Emergence of grouping in multi-resource minority game dynamics," Nature Scientific Reports 2, 703, 1-8 (2012).

29. Y.-Z. Chen and Y.-C. Lai, "Optimizing cooperation on complex networks in the presence of failure," Physical Review E (Rapid Communications) 86, 045101(R), 1-4 (2012).

30. L. Huang, Y.-C. Lai, and M. A. F. Harrison, "Probing complex networks from measured time series," International Journal of Bifurcation and Chaos 22, 1250236, 1-12 (2012).

31. H.-X. Yang, W.-X. Wang, and Y.-C. Lai, "Traffic-driven epidemic outbreak on complex networks: how long does it take?" Chaos 22, 043146, 1-5 (2012).

32. Z. Zhou, Z.-G. Huang, L. Huang, Y.-C. Lai, L. Yang, and D.-S. Xue, "Universality of flux-fluctuation law in complex dynamical systems," Physical Review E 87, 012808, 1-6 (2013).

33. J.-Q. Zhang, Z.-G. Huang, J.-Q. Dong, L. Huang, and Y.-C. Lai, "Controlling collective dynamics in complex, minority-game resource-allocation systems," Physical Review E 87, 052808, 1-9 (2013).

34. J.-P. Park, Y.-H. Do, Z.-G. Huang, and Y.-C. Lai, "Persistent coexistence of cyclically competing species in spatially extended ecosystems," Chaos 23, 023128, 1-9 (2013).

35. Z.-Z. Yuan, C. Zhao, Z.-R. Di, W.-X. Wang, and Y.-C. Lai, "Exact controllability of complex net- works," Nature Communications 4, 2447, 1-9 (2013).

36. B.-S. Kim, Y.-H. Do, and Y.-C. Lai, "Emergence of synchronization and size scaling in moving-agent networks," Physical Review E 88, 042818, 1-7 (2013).

37. Z.-D. Zhao, Z.-M. Yang, Z.-K. Zhang, T. Zhou, Z.-G. Huang, and Y.-C. Lai, "Emergence of scaling in human-interest dynamics," Nature Scientific Reports 3, 3472, 1-7 (2013).

38. K. Gong, M. Tang, P. M. Hui, Y. Do, and Y.-C. Lai, "An efficient immunization strategy for community networks," PLoS One 8, e83489, 1-11 (2013).

39. N. Yao, Z.-G. Huang, Y.-C. Lai, and Z.-G. Zheng, "Robustness of chimera states in complex dynamical systems," Nature Scientific Reports 3, 3522, 1-8 (2013).

40. R.-Q. Su, Y.-C. Lai, X. Wang, and Y.-H. Do, "Uncovering hidden nodes in complex network in the presence of noise," Nature Scientific Reports 4, 3944, 1-7 (2014).

41. W. Wang, M. Tang, H. Yang, Y.-H. Do, Y.-C. Lai, and G.-W. Lee, “Asymmetrically interacting spreading dynamics on complex layered networks,” Nature Scientific Reports 4, 5097, 1-8 (2014).

42. Z.-S. Shen, W.-X. Wang, Y. Fan, Z.-R. Di, and Y.-C. Lai, “Reconstructing propagation networks with natural diversity and identifying hidden source,” Nature Communications 5, 4323, 1-10 (2014).

43. H.-F. Zhang, Z.-X. Wu, M. Tang, and Y.-C. Lai, “Effects of behavioral response and vaccination policy on epidemic spreading - an approach based on evolutionary-game dynamics,” Nature Scientific Reports 4, 5666, 1-10 (2014).

44. R.-Q. Su, Y.-C. Lai, and X. Wang, “Identifying chaotic FitzHugh-Nagumo neurons using compressive sensing,” Entropy 16, 3889-3902 (2014).

45. Y.-C. Lai, “Controlling complex, nonlinear dynamical networks,” National Science Review 1, 339-341 (2014).

46. Y.-Z. Chen, Z.-G. Huang, and Y.-C. Lai, “Controlling extreme events on complex networks,” Nature Scientific Reports 4, 6121, 1-10 (2014).

47. H.-F. Zhang, J.-R. Xie, M. Tang, and Y.-C. Lai, “Suppression of epidemic spreading in complex networks by local information based behavioral responses,” Chaos 24, 043106, 1-7 (2014).

48. Z.-Z. Yuan, C. Zhao, W.-X. Wang, Z.-R. Di, and Y.-C. Lai, “Exact controllability of multiplex net- works,” New Journal of Physics 16, 103036, 1-24 (2014).

49. L.-Z. Wang, Z.-G. Huang, Z.-H. Rong, X.-F. Wang, and Y.-C. Lai, “Emergence and evolution of online social networks,” PLoS ONE 9(11), e111013, 1-6 (2014).

50. Y.-Z. Chen, Z.-G. Huang, S.-H. Xu, and Y.-C. Lai, “Spatiotemporal patterns and predictability of cyberattacks,” PLoS ONE, accepted (to appear in June 2015).

Changes in research objectives (if any):

None

Change in AFOSR Program Manager, if any:

Original AFOSR Program Manager:

Dr. Robert L. Herklotz
Program Manager: Information Operations and Security
Air Force Office of Scientific Research (AFOSR/RSL)
Suite 325, Room 3112
875 N. Randolph Street
Arlington, VA 22203-1768

Current AFOSR Program Manager:

Dr. Tristan Nguyen
Program Manager: Information Operations and Security
Air Force Office of Scientific Research (AFOSR/RSL)
Suite 325, Room 3112
875 N. Randolph Street
Arlington, VA 22203-1768

Extensions granted or milestones slipped, if any:

None

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

May 29, 2015 13:13:30 Success: Email Sent to: Ying-Cheng.Lai@asu.edu